

# Nash and Wardrop equilibria in aggregative games with coupling constraints

Basilio Gentile, Francesca Parise, Dario Paccagnan, Maryam Kamgarpour and John Lygeros

**Abstract**—We consider the framework of aggregative games, in which the cost function of each agent depends on his own strategy and on the average population strategy. As first contribution, we investigate the relations between the concepts of Nash and Wardrop equilibrium. By exploiting a characterization of the two equilibria as solutions of variational inequalities, we bound their distance with a decreasing function of the population size. As second contribution, we propose two decentralized algorithms that converge to such equilibria and are capable of coping with constraints coupling the strategies of different agents. Finally, we study the applications of charging of electric vehicles and of route choice on a road network.

## I. INTRODUCTION

COMPLEX systems resulting from the interconnection of selfish agents have attracted an increasing interest in the scientific community over the last decade for their ubiquitous appearance in real-life applications and the inherent mathematical challenges that they present. Among the vast literature of non-cooperative game theory, *aggregative games* [1] describe systems where each agent is not subject to a one-to-one interaction, but is rather influenced by an aggregate quantity depending on the strategies of the entire population. The vast spectrum of their applications ranges from traffic [2] or transmission networks [3] to electricity [4] or commodity markets [5]. Extending our preliminary work [6], in the rest of the paper we focus on aggregative games where the aggregate quantity is the average population strategy.

### *Nash and Wardrop equilibria*

A fundamental concept in game theory is the notion of *Nash equilibrium*, which is a set of strategies where no agent can improve his cost by unilaterally changing his strategy. Note that in aggregative games an agent can indirectly influence his cost through his contribution to the average strategy. However, when the population becomes large such contribution becomes negligible. This consideration motivates the introduction of the *Wardrop equilibrium*, which describes a configuration where no agent can improve his cost by changing his strategy, under the assumption that he has no influence on the average. Depending on the application, the literature on aggregative games focuses on either Nash (see [5], [7] in economics, [3] in communication networks) or Wardrop equilibria (see [8]

in network congestion games, [9] in road networks, [10] in electricity markets), with the underlying assumption that the two are close to each other in large populations. In some cases it is possible to achieve one equilibrium but not the other, hence establishing bounds on the distance between them can be useful to quantify the accuracy when approximating the desired equilibrium with the other one. As first contribution of the paper we leverage on the theory of variational inequality [11] to

- present a unifying framework to characterize both Nash and Wardrop equilibria for aggregative games,
- formalize the intuition that in large aggregative games the two equilibria are close by bounding their distance with a decreasing function of the population size.

To the best of our knowledge, the only contributions on the distance between the two equilibria are the results of [8], which focuses on road network games to show that the two equilibria converge to each other, under the restrictive assumption that the population increases by means of identical replicas of the agents. Here we relax this assumption by introducing arbitrary new agents. It is important to note that the recently developed theory of mean field games [12], [13] shows in a stochastic and unconstrained setup that a Wardrop equilibrium is an  $\varepsilon$ -Nash equilibrium, but the distance between the equilibrium strategies is not investigated.

### *Decentralized algorithms and coupling constraints*

The second part of the paper focuses on coordinating the strategies of the agents to a Nash or a Wardrop equilibrium. For reasons of privacy and computational intractability of centralized solutions, we focus on decentralized algorithms, based on the presence of a central operator capable of iteratively broadcasting a common signal to the population. Following the large literature on this topic, we consider two different scenarios based on whether the agents respond to the common signal by solving a minimization problem (*optimal response*) as in [14], [15] or by taking a *gradient step* as in [16], [17]. Our algorithms differ from all the aforementioned works in that we handle constraints coupling the agents' decisions. Specifically, building upon [18], we contribute as follows:

- we propose a decentralized two-level algorithm based on optimal response, which integrates the scheme proposed in [14] with an outer loop that updates a dual variable to achieve a Wardrop equilibrium;
- we propose a decentralized one-level asymmetric projection algorithm based on gradient step to achieve either a Nash or a Wardrop equilibrium.

While coupling constraints are of fundamental importance in many applications, such as electricity markets [19], or

Date: March 1, 2017. The first three authors contributed equally. This work was supported by the European Commission project DYMASOS (FP7-ICT 611281), by the Swiss Competence Centers for Energy Research FEEB&D, by the ERC Starting Grant CONENE and by the SNSF grant number P2EZP2\_168812. B. Gentile, D. Paccagnan, M. Kamgarpour and J. Lygeros are with the Automatic Control Laboratory, ETH Zürich, Switzerland. {gentileb,dariop,mkamgar,lygeros}@control.ee.ethz.ch. F. Parise is with the Laboratory for Information and Decision Systems, MIT, Cambridge, MA, USA. parisef@mit.edu The authors thank Marius Schmitt for fruitful discussion on the traffic model of Section VII and Ricardo Campos for helping with Lemma 5.

communication networks [20], we are not aware of previous schemes that take them into account within the literature of aggregative games. Outside the game theoretical framework, our algorithms have some connection with the ones suggested in [21] for multi-user optimization, where however the agents do not influence the cost functions of the others.

### Applications

*Charging of Electric Vehicles:* Electric-vehicles (EV) are foreseen to significantly penetrate the market in the coming years [22], therefore coordinating their charging schedules can provide services beneficial to the grid operations [23], [24]. By assuming that the electricity price depends on the aggregate consumption, [10], [14], [16] formulate the EV charging problem as an aggregative game and propose decentralized schemes based on optimal response or gradient step, in the absence of coupling constraints. The proposed schemes steer the population to Nash [16] or Wardrop [10], [14] equilibria. We extend the existing literature by introducing constraints coupling the agents' charging profiles. Such constraints model limits on the aggregate peak consumption or on the local consumption of EVs connected to the same transformer. We build upon our theoretical findings to derive ad-hoc results for the EV game. Finally, we establish uniqueness of the dual variables associated to the violation of the coupling constraints.

*Route choice on a road network:* Traffic congestion is a well-recognized issue in modern cities, and the corresponding economic costs are significant [25]. Since every driver seeks his own interest (e.g., minimizing the travel time) and is affected by the others' choices via congestion, a classic approach is to model the traffic problem as a game [26]. Specializing [11, Section 1.4.5], we focus on a stationary model that aims at capturing the basic interactions among the vehicles flow during rush hours. Building upon our theoretical findings, we derive ad-hoc results for the route choice game. Moreover, we perform a realistic numerical analysis based on the data set of the city of Oldenburg in Germany [27]. Specifically, we investigate via simulation the effect of road access limitations, expressed as coupling constraints [28].

*Organization:* Sections II and III introduce game and preliminary results. Sections IV and V present our main contributions, namely the bound on the distance between Nash and Wardrop equilibria and the design of decentralized algorithms to achieve them. Sections VI and VII focus on the applications.

*Notation:*  $\|x\|$  is the 2-norm of  $x \in \mathbb{R}^n$ .  $I_n \in \mathbb{R}^{n \times n}$  is the identity matrix,  $\mathbf{1}_n \in \mathbb{R}^n$  is the vector of unit entries,  $\mathbf{0}_n \in \mathbb{R}^n$  is the vector of zero entries,  $e_i$  is the  $i^{\text{th}}$  canonical vector. Given  $A \in \mathbb{R}^{n \times n}$ ,  $A \succ 0$  ( $\succeq 0$ )  $\Leftrightarrow x^\top A x > 0$  ( $\geq 0$ ),  $\forall x \neq 0$ ;  $\|A\|$  is the induced 2-norm of  $A$ . Given  $M$  vectors each in  $\mathbb{R}^n$ ,  $[x^1; \dots; x^M] := [x^i]_{i=1}^M := [x^1^\top, \dots, x^M^\top]^\top \in \mathbb{R}^{Mn}$  and  $x^{-i} := [x_1; \dots; x_{i-1}; x_{i+1}; \dots; x_M] \in \mathbb{R}^{(M-1)n}$ . Given a matrix  $A \in \mathbb{R}^{m \times Mn}$ ,  $A_{(:,i)} \in \mathbb{R}^{m \times n}$  is such that  $A = [A_{(:,1)}, \dots, A_{(:,M)}]$ . Given  $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  we define  $\nabla_x g(x) \in \mathbb{R}^{n \times m}$  with  $[\nabla_x g(x)]_{i,j} := \frac{\partial g_j(x)}{\partial x^i}$ . Given  $g(x) : \mathbb{R} \rightarrow \mathbb{R}$ , we denote  $g'(x) = \frac{\partial g(x)}{\partial x}$ . Given the sets  $\mathcal{X}^1, \dots, \mathcal{X}^M \subseteq \mathbb{R}^n$ , we denote  $\frac{1}{M} \sum_{i=1}^M \mathcal{X}^i := \{z \in \mathbb{R}^n | z = \frac{1}{M} \sum_{i=1}^M x^i, \text{ for some } x^i \in \mathcal{X}^i\}$ .

## II. PROBLEM FORMULATION

We consider a population of  $M$  agents. Each agent can choose his strategy  $x^i$  in his individual constraint set  $\mathcal{X}^i \subset \mathbb{R}^n$ . For finite horizon discrete-time dynamic games the constraint set can represent the dynamics of a system [10], [14]. We assume that the cost function  $J^i(x^i, \sigma(x))$  of agent  $i$  depends on his own strategy  $x^i \in \mathcal{X}^i$  and on the strategies of the other agents via the average population strategy  $\sigma(x) := \frac{1}{M} \sum_{j=1}^M x^j \in \frac{1}{M} \sum_{i=1}^M \mathcal{X}^i$ , as typical of aggregative games [1]. The expression of the cost function of agent  $i$  is

$$J^i(x^i, \sigma(x)) := v^i(x^i) + p(\sigma(x))^\top x^i. \quad (1)$$

The cost in (1) can for example describe applications where  $x^i$  denotes the usage level of a certain commodity, whose negative utility is modeled by  $v^i : \mathcal{X}^i \rightarrow \mathbb{R}$  and whose per-unit cost  $p : \frac{1}{M} \sum_{i=1}^M \mathcal{X}^i \rightarrow \mathbb{R}^n$  depends on the average usage level of the entire population [4], [10]. Besides the individual constraints, each agent has to satisfy a linear coupling constraint, which involves the decision variables of other agents. Upon defining  $x = [x^1; \dots; x^M] \in \mathbb{R}^{Mn}$ , the coupling constraint can be expressed as

$$x \in \mathcal{C} := \{x \in \mathbb{R}^{Mn} | Ax \leq b\} \subset \mathbb{R}^{Mn}, \quad (2)$$

with  $A := [A_{(:,1)}, \dots, A_{(:,M)}] \in \mathbb{R}^{m \times Mn}$ ,  $A_{(:,i)} \in \mathbb{R}^{m \times n}$  for all  $i \in \{1, \dots, M\}$ ,  $b \in \mathbb{R}^m$ . The coupling constraint in (2) can model for instance the fact that the overall usage level for a certain commodity cannot exceed a fixed capacity. The cost and constraints just introduced give rise to the game

$$\mathcal{G} := \begin{cases} \text{agents :} & \{1, \dots, M\} \\ \text{cost of agent } i : & J^i(x^i, \sigma(x)) \\ \text{individual constraint :} & \mathcal{X}^i \\ \text{coupling constraint :} & \mathcal{C}, \end{cases} \quad (3)$$

which is the focus of the rest of the paper. We denote for convenience  $\mathcal{X} := \mathcal{X}^1 \times \dots \times \mathcal{X}^M$  and define

$$\mathcal{Q}^i(x^{-i}) := \{x^i \in \mathcal{X}^i | Ax \leq b\}, \quad \mathcal{Q} := \mathcal{X} \cap \mathcal{C}. \quad (4)$$

### A. Equilibrium definitions

We consider two notions of equilibrium for the game  $\mathcal{G}$  in (3). The first is a known generalization of the concept of Nash equilibrium to games with coupling constraints [18].

**Definition 1** (Nash Equilibrium). *A set of strategies  $x_N = [x_N^1; \dots; x_N^M] \in \mathbb{R}^{Mn}$  is an  $\varepsilon$ -Nash equilibrium of the game  $\mathcal{G}$  if  $x_N \in \mathcal{Q}$  and for all  $i \in \{1, \dots, M\}$  and all  $x^i \in \mathcal{Q}^i(x_N^{-i})$*

$$J^i(x_N^i, \sigma(x_N)) \leq J^i\left(x^i, \frac{1}{M} x^i + \frac{1}{M} \sum_{j \neq i} x_N^j\right) + \varepsilon. \quad (5)$$

If (5) holds with  $\varepsilon = 0$  then  $x_N$  is a Nash equilibrium.  $\square$

Intuitively, a feasible set of strategies  $\{x_N^i\}_{i=1}^M$  is a Nash equilibrium if no agent can improve his cost by unilaterally deviating from his strategy, assuming that the strategies of the other agents are fixed. A Nash equilibrium for a game with coupling constraints is usually referred to as generalized Nash equilibrium [18]; in this paper we omit the word generalized, even though we consider a game with coupling constraints.

Note that on the right-hand side of (5) the decision variable  $x^i$  appears in both arguments of  $J^i(\cdot, \cdot)$ . However, as the population size grows the contribution of agent  $i$  to the average  $\sigma(x)$  decreases. This observation motivates the definition of Wardrop equilibrium.

**Definition 2** (Wardrop Equilibrium). *A set of strategies  $x_W = [x_W^1; \dots; x_W^M] \in \mathbb{R}^{Mn}$  is a Wardrop equilibrium of the game  $\mathcal{G}$  if  $x_W \in \mathcal{Q}$  and for all  $i \in \{1, \dots, M\}$  and all  $x^i \in \mathcal{Q}^i(x_W^{-i})$*

$$J^i(x_W^i, \sigma(x_W)) \leq J^i(x^i, \sigma(x_W)). \quad \square$$

Intuitively, a feasible set of strategies  $\{x_W^i\}_{i=1}^M$  is a Wardrop equilibrium if no agent can improve his cost by unilaterally deviating from his strategy, assuming that the average strategy is fixed. We note that Definition 2 generalizes to aggregative games with coupling constraints the notions of Wardrop equilibrium introduced in [26] for congestion games.

### III. CONNECTION WITH VARIATIONAL INEQUALITIES

This section shows that some equilibria of the game  $\mathcal{G}$  in (3) can be obtained by solving a variational inequality. This fact is then used to derive the results of Sections IV and V.

**Definition 3** (Variational inequality [11]). *Consider a set  $\mathcal{K} \subseteq \mathbb{R}^d$  and an operator  $F : \mathcal{K} \rightarrow \mathbb{R}^d$ . A point  $\bar{x} \in \mathcal{K}$  is a solution of the variational inequality  $\text{VI}(\mathcal{K}, F)$  if*

$$F(\bar{x})^\top (x - \bar{x}) \geq 0, \quad \forall x \in \mathcal{K}. \quad \square$$

Let us define

$$F_N(x) := [\nabla_{x^i} J^i(x^i, \sigma(x))]_{i=1}^M, \quad (6a)$$

$$F_W(x) := [\nabla_{x^i} J^i(x^i, z)]_{i=1}^M \Big|_{z=\sigma(x)}, \quad (6b)$$

where  $F_N(x), F_W(x) : \mathcal{X} \rightarrow \mathbb{R}^{Mn}$ . The operator  $F_N$  is obtained by stacking together the gradients of each agent's cost with respect to his decision variable.  $F_W$  is obtained similarly, but considering  $\sigma(x)$  as fixed when differentiating. The following proposition provides a sufficient characterization of the equilibria described in Definitions 1 and 2 as solutions of two variational inequalities, which feature the same set  $\mathcal{Q}$ , defined in (4), but different operators, namely  $F_N$  and  $F_W$  in (6).

**Assumption 1.** *For all  $i \in \{1, \dots, M\}$ , the individual constraint set  $\mathcal{X}^i$  is closed and convex. The set  $\mathcal{Q}$  in (4) is non-empty. The cost functions  $J^i(x^i, \sigma(x))$  are convex and continuously differentiable in  $x^i$  for any fixed  $\{x^j \in \mathcal{X}^j\}_{j \neq i}$ . The cost functions  $J^i(x^i, z)$  are convex and continuously differentiable in  $x^i$  for any  $z \in \frac{1}{M} \sum_{i=1}^M \mathcal{X}^i$ .*  $\square$

**Proposition 1.** *Under Assumption 1, the following hold.*

- 1) Any solution  $\bar{x}_N$  of  $\text{VI}(\mathcal{Q}, F_N)$  is a Nash equilibrium of the game  $\mathcal{G}$  in (3);
- 2) Any solution  $\bar{x}_W$  of  $\text{VI}(\mathcal{Q}, F_W)$  is a Wardrop equilibrium of the game  $\mathcal{G}$  in (3).  $\square$

*Proof:* The proof of the first statement can be found in [29, Theorem 2.1], we prove the second one. We rewrite the operator  $F_W(x)$  as  $\tilde{F}_W(x, \sigma(x))$ , where  $\tilde{F}_W(x, z) := [\nabla_{x^i} J^i(x^i, z)]_{i=1}^M$ .

By definition, if  $\bar{x}_W$  solves  $\text{VI}(\mathcal{Q}, F_W)$  then  $F_W(\bar{x}_W)^\top (x - \bar{x}_W) \geq 0$  for all  $x \in \mathcal{Q}$ , i.e.

$$\tilde{F}_W(\bar{x}_W, \bar{z})^\top (x - \bar{x}_W) \geq 0, \quad \forall x \in \mathcal{Q}, \quad \bar{z} = \sigma(\bar{x}_W). \quad (7)$$

Consider  $i \in \{1, \dots, M\}$ , set  $x^{-i} = \bar{x}_W^{-i}$  in (7) and consider an arbitrary  $x^i \in \mathcal{Q}^i(\bar{x}_W^{-i})$ ; then all the summands in (7) vanish except the  $i^{\text{th}}$  one and (7) reads

$$\nabla_{x^i} J^i(\bar{x}_W^i, \bar{z})^\top (x^i - \bar{x}_W^i) \geq 0, \quad \forall x^i \in \mathcal{Q}^i(\bar{x}_W^{-i}). \quad (8)$$

Consider the convex function  $J^i(\cdot, \bar{z}) : \mathcal{Q}^i(\bar{x}_W^{-i}) \rightarrow \mathbb{R}$ . Since  $\mathcal{Q}^i(\bar{x}_W^{-i})$  is a convex set, by (8) and [30, Proposition 3.1] we have that  $\bar{x}_W^i \in \arg\min_{x^i \in \mathcal{Q}^i(\bar{x}_W^{-i})} J^i(x^i, \bar{z})$ . Substituting  $\bar{z} = \sigma(\bar{x}_W)$ , one has  $J^i(\bar{x}_W^i, \sigma(\bar{x}_W)) \leq J^i(x^i, \sigma(\bar{x}_W))$  for all  $x^i \in \mathcal{Q}^i(\bar{x}_W^{-i})$ . Since this holds for all  $i \in \{1, \dots, M\}$  and since  $\bar{x}_W \in \mathcal{Q}$ , it follows that  $\bar{x}_W$  is a Wardrop equilibrium of  $\mathcal{G}$ .  $\blacksquare$

Proposition 1 states that a solution of the variational inequality is an equilibrium. The converse in general does not hold due to the presence of the coupling constraints. If on the other hand  $\mathcal{C} = \mathbb{R}^{Mn}$ , then  $\mathcal{Q} = \mathcal{X}$  and one can show that  $x_N$  solves the  $\text{VI}(\mathcal{Q}, F_N)$  if and only if it is a Nash equilibrium of  $\mathcal{G}$  and  $x_W$  solves the  $\text{VI}(\mathcal{Q}, F_W)$  if and only if it is a Wardrop equilibrium of  $\mathcal{G}$  [18, Corollary 1]. The equilibria that can be obtained as solution of the corresponding variational inequality are called *variational equilibria* [18, Definition 3] and are here denoted with  $\bar{x}_N, \bar{x}_W$  instead of  $x_N, x_W$ , that indicate any equilibria satisfying Definitions 1 and 2.

In the following we provide sufficient conditions for existence and uniqueness of variational equilibria.

**Definition 4** (Strong monotonicity [11]). *An operator  $F : \mathcal{K} \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$  is strongly monotone on the set  $\hat{\mathcal{K}} \subseteq \mathcal{K}$  with monotonicity constant  $\alpha > 0$  if<sup>1</sup>*

$$(F(x) - F(y))^\top (x - y) \geq \alpha \|x - y\|^2, \quad \forall x, y \in \hat{\mathcal{K}}. \quad (9)$$

*The operator is monotone on  $\hat{\mathcal{K}}$  if (9) holds for  $\alpha = 0$ .*  $\square$

**Lemma 1.** [11, Corollary 2.2.5, Theorem 2.3.3] *Let Assumption 1 hold. Then*

- 1) *If  $\mathcal{X}^i$  is bounded for all  $i \in \{1, \dots, M\}$ , then both  $\text{VI}(\mathcal{Q}, F_N)$  and  $\text{VI}(\mathcal{Q}, F_W)$  admit a solution<sup>2</sup>.*
- 2) *If  $F_N$  is strongly monotone on  $\mathcal{Q}$ , then  $\text{VI}(\mathcal{Q}, F_N)$  has a unique solution. If  $F_W$  is strongly monotone on  $\mathcal{Q}$  then  $\text{VI}(\mathcal{Q}, F_W)$  has a unique solution.*  $\square$

To verify whether an operator is strongly monotone or monotone one can exploit the following equivalent characterizations.

**Lemma 2.** [11, Proposition 2.3.2] *A continuously differentiable operator  $F : \mathcal{K} \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$  is strongly monotone with monotonicity constant  $\alpha$  (resp. monotone) if and only if  $\nabla_x F(x) \succeq \alpha I$  (resp.  $\nabla_x F(x) \succeq 0$ ) for all  $x \in \mathcal{K}$ . Moreover, if  $\mathcal{X}$  is compact then there exists  $\alpha > 0$  such that  $\nabla_x F(x) \succeq \alpha I$  for all  $x \in \mathcal{K}$  if and only if  $\nabla_x F(x) \succ 0$  for all  $x \in \mathcal{K}$ .*  $\square$

<sup>1</sup>When we do not specify the set  $\hat{\mathcal{K}}$  this is understood to be  $\mathcal{K}$ , i.e. the domain of the operator.

<sup>2</sup>The convexity of the cost functions required by Assumption 1 is not needed for the first statement of Lemma 1, continuity is enough.

The previous lemma can be used to derive sufficient conditions for strong monotonicity of the operators  $F_N$  and  $F_W$ . To this end, we specialize (6) to the cost function given in (1)

$$F_W(x) = [\nabla_{x^i} v^i(x^i)]_{i=1}^M + [p(\sigma(x))]_{i=1}^M, \quad (10a)$$

$$F_N(x) = F_W(x) + \frac{1}{M} [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M. \quad (10b)$$

**Lemma 3.**

- 1) Suppose that for each agent  $i \in \{1, \dots, M\}$  the function  $v^i$  in (1) is convex and that  $p$  in (1) is monotone; then  $F_W$  is monotone. Under the further assumption that  $p$  is affine and strongly monotone,  $F_N$  is strongly monotone.
- 2) Suppose that for each agent  $i \in \{1, \dots, M\}$  the function  $v^i$  is strongly convex and that  $p$  is monotone. Then  $F_W$  is strongly monotone.  $\square$

The proof is given in the appendix. It is clear from Lemma 3 that often only one of  $F_N$  and  $F_W$  possesses monotonicity properties, which are required to guarantee that an equilibrium can be achieved using the algorithms proposed in Section V. Hence it is important to derive results on the distance between the two equilibria, which is the goal of the next Section IV.

#### IV. DISTANCE BETWEEN NASH AND WARDROP EQUILIBRIA IN LARGE POPULATIONS

In this section we study the relations between Nash and Wardrop equilibria in aggregative games with large populations. Specifically, we consider a sequence of games  $(\mathcal{G}_M)_{M=1}^\infty$ . For fixed  $M$ , the game  $\mathcal{G}_M$  is played among  $M$  agents and is defined as in (3) with an arbitrary coupling constraint  $\mathcal{C}$  and, for every agent  $i$ , arbitrary  $v^i$  and  $\mathcal{X}^i$ . For the sake of readability, we avoid the explicit dependence on  $M$  in denoting these quantities and in denoting  $x_N, x_W, F_N, F_W$ . The function  $p$  is instead the same for every game of the sequence  $(\mathcal{G}_M)_{M=1}^\infty$ .

**Assumption 2.** There exists a convex, compact set  $\mathcal{X}^0 \subset \mathbb{R}^n$  such that  $\cup_{i=1}^M \mathcal{X}^i \subseteq \mathcal{X}^0$  for each  $\mathcal{G}_M$  in the sequence  $(\mathcal{G}_M)_{M=1}^\infty$ . The function  $p$  is Lipschitz on the set  $\mathcal{X}^0$  with Lipschitz constant  $L_p$ .  $\square$

We note that Assumption 2 implies that  $\sigma(x) \in \mathcal{X}^0$  for any  $M$  and any  $x \in \mathcal{X}^1 \times \dots \times \mathcal{X}^M$ . Moreover, under Assumption 2 we define  $R := \max_{y \in \mathcal{X}^0} \{\|y\|\}$ . The next proposition shows that every Wardrop equilibrium is an  $\varepsilon$ -Nash equilibrium, with  $\varepsilon$  tending to zero as  $M$  grows.

**Proposition 2.** Let the sequence of games  $(\mathcal{G}_M)_{M=1}^\infty$  satisfy Assumption 2. For each  $\mathcal{G}_M$ , every Wardrop equilibrium is an  $\varepsilon$ -Nash equilibrium, with  $\varepsilon = \frac{2R^2 L_p}{M}$ .  $\square$

*Proof:* Consider any Wardrop equilibrium  $x_W$  of  $\mathcal{G}_M$  (not necessarily a variational one). By Definition 2,  $x_W \in \mathcal{Q}$  and for each agent  $i$

$$J^i(x_W^i, \sigma(x_W)) \leq J^i(x^i, \sigma(x_W)), \quad \text{for all } x^i \in \mathcal{Q}^i(x_W^{-i}).$$

It follows that for each agent  $i$  and for all  $x^i \in \mathcal{Q}^i(x_W^{-i})$

$$\begin{aligned} & J^i(x_W^i, \sigma(x_W)) - J^i(x^i, \frac{1}{M}(x^i + \sum_{j \neq i} x_W^j)) \\ &= \underbrace{J^i(x_W^i, \sigma(x_W)) - J^i(x^i, \sigma(x_W))}_{\leq 0} + \\ & J^i(x^i, \sigma(x_W)) - J^i(x^i, \frac{1}{M}(x^i + \sum_{j \neq i} x_W^j)) \\ &\leq |p(\sigma(x_W))x^i - p(\frac{1}{M}(x^i + \sum_{j \neq i} x_W^j))x^i| \\ &\leq \|p(\sigma(x_W)) - p(\frac{1}{M}(x^i + \sum_{j \neq i} x_W^j))\| \|x^i\| \\ &\leq \frac{L_p R}{M} \|(x_W^i + \sum_{j \neq i} x_W^j) - (x^i + \sum_{j \neq i} x_W^j)\| \\ &\leq \frac{L_p R}{M} \|x_W^i - x^i\| \leq \frac{2R^2 L_p}{M}, \end{aligned}$$

hence  $x_W$  is an  $\varepsilon$ -Nash equilibrium of  $\mathcal{G}_M$ . This result was shown in [14, Theorem 1] for a game without coupling constraints and with a different proof.  $\blacksquare$

Proposition 2 is a strong result but it provides no information on the distance between the set of strategies constituting a Nash and the set of strategies constituting a Wardrop equilibrium. In the following we study this distance for variational equilibria.

**Theorem 1.** Let the sequence of games  $(\mathcal{G}_M)_{M=1}^\infty$  satisfy Assumption 2, and each  $\mathcal{G}_M$  satisfy Assumption 1. Then the following hold.

- 1) If the operator  $F_N$  relative to  $\mathcal{G}_M$  is strongly monotone on  $\mathcal{Q}$  with monotonicity constant  $\alpha_M > 0$ , then there exists a unique variational Nash equilibrium  $\bar{x}_N$  of  $\mathcal{G}_M$ . Moreover, for any variational Wardrop equilibrium  $\bar{x}_W$

$$\|\bar{x}_N - \bar{x}_W\| \leq \frac{L_p R}{\alpha_M \sqrt{M}}. \quad (11)$$

As a consequence, if  $\alpha_M \sqrt{M} \rightarrow \infty$  as  $M \rightarrow \infty$ , then  $\|\bar{x}_N - \bar{x}_W\| \rightarrow 0$  as  $M \rightarrow \infty$ .

- 2) If the operator  $F_W$  relative to  $\mathcal{G}_M$  is strongly monotone on  $\mathcal{Q}$  with monotonicity constant  $\alpha_M > 0$ , then there exists a unique variational Wardrop equilibrium  $\bar{x}_W$  of  $\mathcal{G}_M$ . Moreover, for any variational Nash equilibrium  $\bar{x}_N$

$$\|\bar{x}_N - \bar{x}_W\| \leq \frac{L_p R}{\alpha_M \sqrt{M}}. \quad (12)$$

As a consequence, if  $\alpha_M \sqrt{M} \rightarrow \infty$  as  $M \rightarrow \infty$ , then  $\|\bar{x}_N - \bar{x}_W\| \rightarrow 0$  as  $M \rightarrow \infty$ .

- 3) If in the game  $\mathcal{G}_M$  it holds  $v^i = 0$  for all  $i$  and  $p$  is strongly monotone on  $\mathcal{X}^0$  with monotonicity constant  $\alpha$ , then there exists a unique  $\bar{\sigma}$  such that  $\sigma(\bar{x}_W) = \bar{\sigma}$  for any variational Wardrop equilibrium  $\bar{x}_W$  of  $\mathcal{G}_M$ . Moreover, for any variational Nash equilibrium  $\bar{x}_N$  of  $\mathcal{G}_M$  and for any variational Wardrop equilibrium  $\bar{x}_W$  of  $\mathcal{G}_M$

$$\|\sigma(\bar{x}_N) - \sigma(\bar{x}_W)\| \leq \sqrt{\frac{2R^2 L_p}{\alpha M}}. \quad (13)$$

Hence,  $\|\sigma(\bar{x}_N) - \sigma(\bar{x}_W)\| \rightarrow 0$  as  $M \rightarrow \infty$ .  $\square$

*Proof:* 1) We first bound the distance between the operators  $F_N$  and  $F_W$  in terms of  $M$ . By (10) it holds

$$\begin{aligned} & \|F_N(x) - F_W(x)\|^2 \\ &= \|\frac{1}{M} [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M\|^2 \\ &= \frac{1}{M^2} \sum_{i=1}^M \|\nabla_z p(z)|_{z=\sigma(x)} x^i\|^2 \\ &\leq \frac{1}{M^2} \sum_{i=1}^M \|\nabla_z p(z)|_{z=\sigma(x)}\|^2 \|x^i\|^2 \leq \frac{L_p^2 R^2}{M}, \end{aligned}$$

where we used  $R := \max_{y \in \mathcal{X}^0} \|y\|$  and the fact that  $p$  is Lipschitz on  $\mathcal{X}^0$  with constant  $L_p$  by Assumption 2, hence  $\|\nabla_z p(z)|_{z=\sigma(x)}\| \leq L_p$ . It follows that

$$\|F_N(x) - F_W(x)\| \leq \frac{L_p R}{\sqrt{M}}. \quad (14)$$

for all  $x \in \mathcal{X}^0$ . We exploit (14) to bound the distance between Nash and Wardrop strategies. Since  $F_N$  is strongly monotone on  $\mathcal{Q}$  by assumption,  $\text{VI}(\mathcal{Q}, F_N)$  has a unique solution  $\bar{x}_N$  by Lemma 1. Moreover, by [31, Theorem 1.14] for all solutions  $\bar{x}_W$  of  $\text{VI}(\mathcal{Q}, F_W)$  it holds

$$\|\bar{x}_N - \bar{x}_W\| \leq \frac{1}{\alpha_M} \|F_N(\bar{x}_W) - F_W(\bar{x}_W)\|.$$

Combining this with equation (14) yields the result.

- 2) As in the above, with Nash in place of Wardrop and viceversa.  
3) Any solution  $\bar{x}_W$  to the  $\text{VI}(\mathcal{Q}, F_W)$  satisfies

$$\begin{aligned} F_W(\bar{x}_W)^\top (x - \bar{x}_W) &\geq 0, \quad \forall x \in \mathcal{Q} \Leftrightarrow \\ \sum_{i=1}^M p(\sigma(\bar{x}_W))^\top (x^i - \bar{x}_W^i) &\geq 0, \quad \forall x \in \mathcal{Q} \Leftrightarrow \\ p(\sigma(\bar{x}_W))^\top (\sigma(x) - \sigma(\bar{x}_W)) &\geq 0, \quad \forall x \in \mathcal{Q}. \end{aligned} \quad (15)$$

Any solution  $\bar{x}_N$  to the  $\text{VI}(\mathcal{Q}, F_N)$  satisfies

$$\begin{aligned} F_N(\bar{x}_N)^\top (x - \bar{x}_N) &\geq 0, \quad \forall x \in \mathcal{Q} \Leftrightarrow \\ p(\sigma(\bar{x}_N))^\top (\sigma(x) - \sigma(\bar{x}_N)) &+ \\ \frac{1}{M^2} \sum_{i=1}^M (\nabla_z p(z)|_{z=\sigma(\bar{x}_N)} \bar{x}_N^i)^\top (x^i - \bar{x}_N^i) &\geq 0, \quad \forall x \in \mathcal{Q}. \end{aligned} \quad (16)$$

Exploiting the strong monotonicity of  $p$  on  $\mathcal{X}^0$ , one has

$$\begin{aligned} &\alpha \|\sigma(\bar{x}_W) - \sigma(\bar{x}_N)\|^2 \\ &\leq (p(\sigma(\bar{x}_W)) - p(\sigma(\bar{x}_N)))^\top (\sigma(\bar{x}_W) - \sigma(\bar{x}_N)) \\ &= p(\sigma(\bar{x}_W))^\top (\sigma(\bar{x}_W) - \sigma(\bar{x}_N)) - p(\sigma(\bar{x}_N))^\top (\sigma(\bar{x}_W) - \sigma(\bar{x}_N)) \\ &\stackrel{\text{by (15)}}{\leq} -p(\sigma(\bar{x}_N))^\top (\sigma(\bar{x}_W) - \sigma(\bar{x}_N)) \\ &\stackrel{\text{by (16)}}{\leq} \frac{1}{M^2} \sum_{i=1}^M (\bar{x}_N^i)^\top (\nabla_z p(z)|_{z=\sigma(\bar{x}_N)})^\top (\bar{x}_W^i - \bar{x}_N^i) \\ &\leq \frac{1}{M^2} \sum_{i=1}^M \|\bar{x}_N^i\| \|\nabla_z p(z)|_{z=\sigma(\bar{x}_N)}\| \|\bar{x}_W^i - \bar{x}_N^i\| \\ &\leq \frac{1}{M^2} \sum_{i=1}^M R L_p 2R \leq \frac{1}{M} 2R^2 L_p. \end{aligned}$$

We conclude that  $\|\sigma(\bar{x}_W) - \sigma(\bar{x}_N)\| \leq \sqrt{\frac{2R^2 L_p}{\alpha M}}$ .  $\blacksquare$

We point out that the bounds (11) and (12) can be used to derive a bound on the average strategies similar to (13).

## V. DECENTRALIZED ALGORITHMS

In this section we turn our attention to the design of algorithms that achieve a Nash or a Wardrop equilibrium. Hence we do not consider a sequence of games as in the previous section, but rather focus on the game (3) with fixed population. We assume that agent  $i$  does not wish to disclose information about his utility function  $v^i$  and individual constraint set  $\mathcal{X}^i$  and that he knows his influence on the coupling constraint, that is, the sub-matrix  $A_{(:,i)}$  in (2). Moreover, we assume the presence of a central operator that is able to measure the population average  $\sigma(x)$ , to evaluate the quantity  $Ax - b$  in (2)

and to broadcast aggregate information to the agents. Based on this information structure, in the following we focus on the design of decentralized algorithms to obtain a solution of either  $\text{VI}(\mathcal{Q}, F_N)$  or  $\text{VI}(\mathcal{Q}, F_W)$ . As the techniques are the same for Nash and Wardrop equilibrium, we consider the general problem  $\text{VI}(\mathcal{Q}, F)$ , where  $F$  can be replaced with  $F_N$  or  $F_W$ .

We start by noting that, if the operator  $F$  is integrable and monotone on  $\mathcal{Q}$ , that is, if there exists a convex function  $E(x) : \mathbb{R}^{Mn} \rightarrow \mathbb{R}$  such that  $F(x) = \nabla_x E(x)$  for all  $x \in \mathcal{Q}$ , then  $\text{VI}(\mathcal{Q}, F)$  is equivalent to the convex optimization problem [11, Section 1.3.1]

$$\underset{x \in \mathcal{Q}}{\operatorname{argmin}} E(x). \quad (17)$$

Therefore a solution of  $\text{VI}(\mathcal{Q}, F)$  and thus a variational equilibrium can be found by applying any of the decentralized optimization algorithms available in the literature [30] to problem (17); the decentralized structure arises because each agent can evaluate  $\nabla_{x^i} E(x)$  by knowing only his strategy  $x^i$  and  $\sigma(x)$ . Equivalently, the integrability assumption guarantees that  $\mathcal{G}$  is a *potential game* with potential function  $E(x)$  [32], hence decentralized convergence tools available for potential games can also be employed [33], [34]. An operator  $F$  is integrable in  $\mathcal{Q}$  if and only if  $\nabla_x F(x)$  is symmetric for all  $x \in \mathcal{Q}$  [11, Theorem 1.3.1]. We anticipate that in both applications of Sections VI and VII the Wardrop operator  $F_W$  in (10a) is integrable but the Nash operator  $F_N$  in (10b) is not.

In the following we intend to find a solution of  $\text{VI}(\mathcal{Q}, F)$  when  $F$  is not necessarily integrable, so that these standard methods cannot be applied. To propose decentralized schemes in presence of coupling constraints, we introduce two reformulations of  $\text{VI}(\mathcal{Q}, F)$  in an extended space  $[x; \lambda]$  where  $\lambda$  are the dual variables relative to the coupling constraint  $\mathcal{C}$ . These two reformulations will then be used to propose two alternative algorithms. Specifically, we define for any  $\lambda \in \mathbb{R}_{\geq 0}^m$  the game

$$\mathcal{G}(\lambda) := \begin{cases} \text{agents :} & \{1, \dots, M\} \\ \text{cost of agent } i : & J^i(x^i, \sigma(x)) + \lambda^\top A(:, i) x^i \\ \text{individual constr :} & \mathcal{X}^i \\ \text{coupling constr :} & \mathbb{R}^{Mn}. \end{cases} \quad (18)$$

Moreover, we introduce the extended  $\text{VI}(\mathcal{Y}, T)$  with

$$\mathcal{Y} := \mathcal{X} \times \mathbb{R}_{\geq 0}^m, \quad T(x, \lambda) := \begin{bmatrix} F(x) + A^\top \lambda \\ -(Ax - b) \end{bmatrix}. \quad (19)$$

The following assumption allows us to draw a connection between  $\text{VI}(\mathcal{Q}, F)$ , the game  $\mathcal{G}(\lambda)$  and  $\text{VI}(\mathcal{Y}, T)$ .

**Assumption 3.** For all  $i \in \{1, \dots, M\}$ , the set  $\mathcal{X}^i$  can be expressed as  $\mathcal{X}^i = \{x^i \in \mathbb{R}^n | g^i(x^i) \leq 0\}$ , where  $g^i : \mathbb{R}^n \rightarrow \mathbb{R}^{p_i}$  is continuously differentiable. The set  $\mathcal{Q}$ , which can thus be expressed as  $\mathcal{Q} = \{x \in \mathbb{R}^{Mn} | g^i(x^i) \leq 0, \forall i, Ax \leq b\}$ , satisfies Slater's constraint qualification as by [35, (5.27)].  $\square$

**Proposition 3.** [36, Section 4.3.2] Let Assumptions 1 and 3 hold. The following statements are equivalent.

- 1) The vector  $\bar{x}$  is a solution of  $\text{VI}(\mathcal{Q}, F)$ .
- 2) There exists  $\bar{\lambda} \in \mathbb{R}_{\geq 0}^m$  such that  $\bar{x}$  is a variational equilibrium of  $\mathcal{G}(\bar{\lambda})$  and  $0 \leq \bar{\lambda} \perp b - A\bar{x} \geq 0$ .

3) There exists  $\bar{\lambda} \in \mathbb{R}_{\geq 0}^m$  such that the vector  $[\bar{x}; \bar{\lambda}]$  is a solution of  $\text{VI}(\mathcal{Y}, T)$ .  $\square$

The proof is sketched in the appendix and is based on [36, Section 4.3.2]. In subsection V-A we exploit the equivalence between 1) and 2) to propose a two-level algorithm based on optimal response that converges to a Wardrop equilibrium. In subsection V-B we leverage on the equivalence between 1) and 3) to propose a one-level algorithm based on gradient step that converges to a Nash equilibrium. The same one-level algorithm can be used to obtain a Wardrop equilibrium, by using  $F_W$  instead of  $F_N$ .

#### A. Two-level algorithm based on optimal response for Wardrop equilibrium

Based on the equivalence between 1) and 2) in Proposition 3, we here introduce Algorithm 1 to achieve a Wardrop equilibrium. The algorithm features an outer loop, in which the central operator broadcasts to the population the dual variables  $\lambda_{(k)}$  based on the current constraint violation, and an inner loop, in which the agents update their strategies to the Wardrop equilibrium of the game  $\mathcal{G}(\lambda_{(k)})$ . Since  $\mathcal{G}(\lambda_{(k)})$  is a game *without* coupling constraints, the Wardrop equilibrium can be found via the iterative algorithm proposed in [14, Alorithm 1]. For each agent  $i \in \{1, \dots, M\}$  we define the optimal response to a signal  $z \in \frac{1}{M} \sum_{i=1}^M \mathcal{X}^i$  and dual variables  $\lambda \in \mathbb{R}_{\geq 0}^m$

$$x_{\text{or}}^i(z, \lambda) := \underset{x^i \in \mathcal{X}^i}{\operatorname{argmin}} J^i(x^i, z) + \lambda^\top A(:, i) x^i. \quad (20)$$

---

#### Algorithm 1 for Wardrop equilibrium

---

**Initialization:** Set  $k = 0$ ,  $\tau > 0$ ,  $x_{(0)}^i \in \mathbb{R}^n$ ,  $\lambda_{(0)} \in \mathbb{R}_{\geq 0}^m$ .

**Iterate until convergence:**

1) Strategies are updated to a Wardrop equilibrium of  $\mathcal{G}_{\lambda_{(k)}}$

**Initialization:** Set  $h = 0$ ,  $\tilde{x}_{(0)}^i = x_{(k)}^i$ ,  $z_{(0)} \in \mathbb{R}^n$ .

**Iterate until convergence:**

$$\tilde{x}_{(h+1)}^i \leftarrow x_{\text{or}}^i(z_{(h)}, \lambda_{(k)}), \forall i \quad (21a)$$

$$\tilde{\sigma}_{(h+1)} \leftarrow \frac{1}{M} \sum_{j=1}^M \tilde{x}_{(h+1)}^j \quad (21b)$$

$$z_{(h+1)} \leftarrow \left(1 - \frac{1}{h}\right) z_{(h)} + \frac{1}{h} \tilde{\sigma}_{(h+1)} \quad (21c)$$

$$h \leftarrow h + 1$$

**Upon convergence:**

$$x_{(k+1)} \leftarrow \tilde{x}_{(h)}$$

2) Dual variables are updated

$$\lambda_{(k+1)} \leftarrow \Pi_{\mathbb{R}_{\geq 0}^m} [\lambda_{(k)} - \tau(b - Ax_{(k+1)})] \quad (22)$$

$$k \leftarrow k + 1.$$


---

The inner loop in Algorithm 1 converges to a Wardrop equilibrium of the game  $\mathcal{G}(\lambda_{(k)})$  under the following assumption.

**Assumption 4.** There exists  $L > 0$  such that, for all  $i \in \{1, \dots, M\}$  and  $\lambda \in \mathbb{R}_{\geq 0}^m$ , the mapping  $z \mapsto x_{\text{or}}^i(z, \lambda)$  is single valued and Lipschitz with constant smaller than  $L$ . Moreover, at least one of the following statements holds.

- 1) For each  $i \in \{1, \dots, M\}$  and  $\lambda \in \mathbb{R}_{\geq 0}^m$ , the mapping  $z \mapsto x_{\text{or}}^i(z, \lambda)$  is non-expansive<sup>3</sup>.
- 2) For each  $i \in \{1, \dots, M\}$  and  $\lambda \in \mathbb{R}_{\geq 0}^m$ , the mapping  $z \mapsto z - x_{\text{or}}^i(z, \lambda)$  is strongly monotone.  $\square$

Sufficient conditions for Assumption 4 to hold are in [14, Corollary 1] for  $v^i$  and  $p$  in (1) respectively quadratic and affine.

**Theorem 2.** Suppose that the operator  $F_W$  in (6b) is strongly monotone on  $\mathcal{X}$  with constant  $\alpha$ , that Assumptions 1, 3, 4 hold, and that  $\mathcal{X}^i$  is bounded for all  $i \in \{1, \dots, M\}$ ; set  $\tau < \frac{2\alpha}{\|A\|^2}$  in (22). Then  $x_{(k)}$  in Algorithm 1 converges to a variational Wardrop equilibrium of  $\mathcal{G}$ .  $\square$

The proof is given in the appendix. To the best of our knowledge this is the first two-level algorithm proposed in the literature for Wardrop equilibrium with coupling constraints. We note that, for the case of  $p$  affine, [37] proposes a one-level optimal response algorithm that converges to a pair  $(\bar{x}, \bar{\lambda})$  such that  $\bar{x}$  is a Wardrop equilibrium of the game  $\mathcal{G}(\bar{\lambda})$  satisfying the coupling constraint. We note that however such point is not a Wardrop equilibrium because the complementarity condition  $0 \leq \bar{\lambda} \perp b - A\bar{x} \geq 0$  is not guaranteed. A two-level gradient-step algorithm for Nash equilibrium with coupling constraints has been proposed in [38, Algorithm 2] and in [39, Section 4].

#### B. Asymmetric projection algorithm based on gradient step for Nash and Wardrop equilibrium

We propose here an algorithm to achieve a Nash or a Wardrop equilibrium by making use of the equivalent reformulation of  $\text{VI}(\mathcal{Q}, F)$  as the extended  $\text{VI}(\mathcal{Y}, T)$  given in Proposition 3. Solving  $\text{VI}(\mathcal{Y}, T)$  instead of  $\text{VI}(\mathcal{Q}, F)$  allows the design of a decentralized algorithm, because the set  $\mathcal{Y}$  is the Cartesian product  $\mathcal{X}^1 \times \dots \times \mathcal{X}^M \times \mathbb{R}_{\geq 0}^m$ , and thus the individual constraint sets  $\mathcal{X}^i$  are decoupled.

Algorithm 2 solves  $\text{VI}(\mathcal{Y}, T)$ , where  $T$  is as in (19), with  $F = F_N$ , and hence achieves a Nash equilibrium. If the same algorithm is used with  $F = F_W$  it achieves a Wardrop equilibrium. At every iteration each agent computes his new strategy  $x_{(k+1)}^i$  by taking a gradient step, based on his previous strategy  $x_{(k)}^i$ , the previous average  $\sigma(x_{(k)})$  and the previous dual variables  $\lambda_{(k)}$ . Given the new coupling constraint violation, the central operator updates the price to  $\lambda_{(k+1)}$  and broadcasts it to the agents.

---

#### Algorithm 2 for Nash and Wardrop equilibrium

---

**Initialization:** Set  $k = 0$ ,  $\tau > 0$ ,  $x_{(0)}^i \in \mathbb{R}^n$ ,  $\lambda_{(0)} \in \mathbb{R}_{\geq 0}^m$ .

**Iterate until convergence:**

$$\sigma_{(k)} \leftarrow \frac{1}{M} \sum_{i=1}^M x_{(k)}^i \quad (23a)$$

$$x_{(k+1)}^i \leftarrow \Pi_{\mathcal{X}^i} [x_{(k)}^i - \tau (\nabla_{x^i} J^i(x_{(k)}^i, \sigma(x_{(k)})) + A_{(:, i)}^\top \lambda_{(k)})], \forall i \quad (23b)$$

$$\lambda_{(k+1)} \leftarrow \Pi_{\mathbb{R}_{\geq 0}^m} [\lambda_{(k)} - \tau(b - 2Ax_{(k+1)} + Ax_{(k)})] \quad (23c)$$

$$k \leftarrow k + 1.$$


---

<sup>3</sup>The mapping is non-expansive if  $\|x_{\text{or}}^i(z_1, \lambda) - x_{\text{or}}^i(z_2, \lambda)\| \leq \|z_1 - z_2\|$  for all  $z_1, z_2$ .

**Theorem 3.** *Let Assumptions 1 and 3 hold. Then*

- *Let  $F_N$  in (6a) be strongly monotone on  $\mathcal{X}$  with constant  $\alpha$  and Lipschitz on  $\mathcal{X}$  with constant  $L_F$ . Set  $\tau > 0$  s.t.*

$$\tau^2 < \frac{1}{\|A\|^2}, \quad (24a)$$

$$\tau < \frac{-L_F + \sqrt{L_F^4 + 4\alpha^2\|A\|^2}}{2\alpha\|A\|^2}. \quad (24b)$$

*Then  $x_{(k)}$  in Algorithm 2 converges to a variational Nash equilibrium of  $\mathcal{G}$  in (3).*

- *Let  $F_W$  in (6b) be strongly monotone on  $\mathcal{X}$  with constant  $\alpha$  and Lipschitz on  $\mathcal{X}$  with constant  $L_F$ , then Algorithm 2 with  $\nabla_{x^i} J^i(x_{(k)}^i, z)_{|z=\sigma(x)}$  in place of  $\nabla_{x^i} J^i(x_{(k)}^i, \sigma(x_{(k)}))$  converges to a variational Wardrop equilibrium, if  $\tau$  satisfies (24).*  $\square$

The proof is given in the appendix and is based on the fact that Algorithm 2 is a specific type of asymmetric projection algorithm [11, Algorithm 12.5.1] applied to  $\text{VI}(\mathcal{V}, T)$ . To the best of our knowledge, convergence of Algorithm 2 has not been proven in our setup. A proof for the case in which  $F$  is affine and symmetric is given in [40, Propositions 2 and 4].

### C. Convergence guarantees for quadratic games

In the previous subsections we have proposed two different algorithms. We summarize in Table I the main conditions that guarantee their convergence.

	Nash	Wardrop
optimal response (Algorithm 1)	-	$F_W$ strongly monotone and Assumption 4
gradient step (Algorithm 2)	$F_N$ strongly monotone	$F_W$ strongly monotone

TABLE I: Range of applicability of the presented algorithms, under Assumptions 1 and 2.

To better understand the differences and the range of applicability of the two algorithms we refine the sufficient conditions of Table I to the important class of aggregative games with quadratic cost functions

$$J^i(x^i, \sigma(x)) := \frac{1}{2}(x^i)^\top Q x^i + (C\sigma(x) + c^i)^\top x^i, \quad (25)$$

where  $Q \in \mathbb{R}^{n \times n}$  is symmetric,  $C \in \mathbb{R}^{n \times n}$ ,  $c^i \in \mathbb{R}^n$ . These cost functions have been used in [13], [14], [41]. Since the operators  $F_N, F_W$  defined in (6) are obtained by differentiating quadratic functions, their expression is affine and can be explicitly characterized as

$$F_W(x) = (I_M \otimes Q + \frac{1}{M} \mathbb{1}_M \mathbb{1}_M^\top \otimes C) x + c, \quad (26a)$$

$$F_N(x) = F_W(x) + \frac{1}{M}(I_M \otimes C^\top)x, \quad (26b)$$

where  $c = [c^1; \dots; c^M]$ . The following lemma exploits the characterization (26) to derive sufficient conditions for strong monotonicity of  $F_W, F_N$  and for Assumption 4. These in turn guarantee convergence of Algorithm 1 and 2 as by Table I.

**Lemma 4.** *The following hold.*

- *If  $Q \succ 0$ ,  $C \succeq 0$  or if  $Q \succeq 0$ ,  $C \succ 0$  then  $F_N$  in (26b) is strongly monotone.*

- *If  $Q \succ 0$ ,  $C \succeq 0$  then  $F_W$  in (26a) is strongly monotone.*
- *If  $Q \succ 0$ ,  $C = C^\top \succ 0$  or if  $Q \succ 0$ ,  $Q - C^\top Q^{-1}C \succ 0$  then  $F_W$  in (26a) is strongly monotone and Assumption 4 is satisfied.*  $\square$

*Proof:* By Lemma 2, strong monotonicity of  $F_W$  in (26a) is equivalent to  $\nabla_x F_W(x) = (I_M \otimes Q + \frac{1}{M} \mathbb{1}_M \mathbb{1}_M^\top \otimes C)^\top \succ 0$ , which is independent from  $x$ . In the same way, strong monotonicity of  $F_N$  in (26b) is equivalent to  $(I_M \otimes Q + \frac{1}{M} \mathbb{1}_M \mathbb{1}_M^\top \otimes C)^\top + \frac{1}{M}(I_M \otimes C^\top)^\top \succ 0$ . Building on this, the first two statements are straightforward to prove. Regarding the last statement,  $Q \succ 0$ ,  $C = C^\top \succ 0$  imply  $\nabla_x F_W(x) \succ 0$ . Moreover, by [14, Theorem 2], Assumption 4.2 is satisfied. By using Schur's theorem, it can be shown that  $Q \succ 0$ ,  $Q - C^\top Q^{-1}C \succ 0$  imply  $Q + C \succ 0$ , hence clearly  $\nabla_x F_W(x) \succ 0$ . Finally, by [14, Theorem 2], Assumption 4.1 is satisfied.  $\blacksquare$

## VI. CHARGING OF ELECTRIC VEHICLES

We model the simultaneous charging of a population of electric vehicles (EV) as a game, following the approach of [10], [14], [16]. Compared to the existing work, our main contributions consist in introducing the coupling constraints, finding a Nash and a Wardrop equilibrium even for the case of  $v^i = 0$  in (1), and studying the distance between the aggregate strategies at the Nash and at the Wardrop equilibrium.

### Constraints

We consider a population of  $M$  electric vehicles. The state of charge of vehicle  $i$  at time  $t$  is described by the variable  $s_t^i$ . The time evolution of  $s_t^i$  is specified by the discrete-time system  $s_{t+1}^i = s_t^i + b^i x_t^i$ ,  $t = 1, \dots, n$ , where  $x_t^i$  is the charging control and the parameter  $b^i > 0$  is the charging efficiency. We assume that the charging control cannot take negative values and that at time  $t$  it cannot exceed  $\tilde{x}_t^i \geq 0$ . The final state of charge is constrained to  $s_{n+1}^i \geq \eta^i$ , where  $\eta^i \geq 0$  is the desired state of charge of agent  $i$ . Denoting  $x^i = [x_1^i, \dots, x_n^i]^\top \in \mathbb{R}^n$ , the individual constraint of agent  $i$  can be expressed as

$$x^i \in \mathcal{X}^i := \left\{ x^i \in \mathbb{R}^n \mid \begin{array}{l} 0 \leq x_t^i \leq \tilde{x}_t^i, \quad \forall t = 1, \dots, n \\ \sum_{t=1}^n x_t^i \geq \theta^i \end{array} \right\}, \quad (27)$$

where  $\theta^i := (b^i)^{-1}(\eta^i - s_1^i)$ , with  $s_1^i \geq 0$  the state of charge at the beginning of the time horizon. Besides the individual constraints  $x^i \in \mathcal{X}^i$ , we also introduce the coupling constraint

$$x \in \mathcal{C} := \{x \in \mathbb{R}^{Mn} \mid \frac{1}{M} \sum_{i=1}^M x_t^i \leq K_t, \quad \forall t = 1, \dots, n\}, \quad (28)$$

indicating that at time  $t$  the grid cannot deliver more than  $M \cdot K_t$  units of power to the vehicles. In compact form (28) reads as  $(\mathbb{1}_M^\top \otimes I_n)x \leq MK$ , where  $K := [K_1, \dots, K_n]^\top$ .

### Cost function

The cost function of each vehicle represents its electricity bill, which we model as

$$J^i(x^i, \sigma(x)) = \sum_{t=1}^n p_t \left( \frac{d_t + \sigma_t(x)}{\kappa_t} \right) x_t^i =: p(\sigma(x))^\top x^i, \quad (29)$$

where we assumed that the energy price for each time interval  $p_t : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{>0}$  depends on the ratio between total consumption and total capacity  $(d_t + \sigma_t(x))/\kappa_t$ , where  $d_t$  and  $\sigma_t(x) := \frac{1}{M} \sum_{i=1}^M x_t^i$  are the non-EV and EV demand at time  $t$  divided by  $M$  and  $\kappa_t$  is the total production capacity divided by  $M$  as in [10, eq. (6)].  $\kappa_t$  is in general not related to  $K_t$ .

#### A. Theoretical guarantees

We define the game  $\mathcal{G}_M^{\text{EV}}$  as in (3), with  $\mathcal{X}^i$ ,  $\mathcal{C}$  and  $J^i(x^i, \sigma(x))$  as in (27), (28) and (29) respectively. In the following corollary we refine the main results of Sections III, IV, V for the EV application.

**Corollary 1.** *Consider a sequence of games  $(\mathcal{G}_M^{\text{EV}})_{M=1}^\infty$ . Assume that there exists  $\tilde{x}^0$  such that  $\tilde{x}_t^i \leq \tilde{x}^0$  for all  $t \in \{1, \dots, n\}$ ,  $i \in \{1, \dots, M\}$  and for each game  $\mathcal{G}_M^{\text{EV}}$ . Moreover, assume that for each game  $\mathcal{G}_M^{\text{EV}}$  the set  $\mathcal{Q} = \mathcal{C} \cap \mathcal{X}$  is non-empty and that for each  $t$  the price function  $p_t$  in (29) is twice continuously differentiable, strictly increasing and Lipschitz in  $[0, \tilde{x}^0]$  with constant  $L_p$ . Moreover, assume*

$$\min_{\substack{t \in \{1, \dots, n\} \\ z \in [0, \tilde{x}^0]}} \left( p'_t(z) - \frac{\tilde{x}^0 p''_t(z)}{8} \right) > 0. \quad (30)$$

Then:

- 1) A Wardrop and a Nash equilibrium exist for each game  $\mathcal{G}_M^{\text{EV}}$  of the sequence. Furthermore, every Wardrop equilibrium is an  $\varepsilon$ -Nash equilibrium with  $\varepsilon = \frac{2n(\tilde{x}^0)^2 L_p}{M}$ .
- 2) The function  $p$  is strongly monotone, hence for each game  $\mathcal{G}_M^{\text{EV}}$  there exists a unique  $\bar{\sigma}$  such that  $\sigma(\bar{x}_W) = \bar{\sigma}$  for any variational Wardrop equilibrium  $\bar{x}_W$  of  $\mathcal{G}_M^{\text{EV}}$ . Moreover for any variational Nash equilibrium  $\bar{x}_N$  of  $\mathcal{G}_M^{\text{EV}}$ ,  $\|\sigma(\bar{x}_N) - \sigma(\bar{x}_W)\| \leq \tilde{x}^0 \sqrt{\frac{2nL_p}{\alpha M}}$ , where  $\alpha$  is the monotonicity constant of  $p$ .
- 3) For each game  $\mathcal{G}_M^{\text{EV}}$  the operator  $F_W$  is monotone, hence the extragradient algorithm [11, Algorithm 12.1.9] with operator  $F_W$  converges to a variational Wardrop equilibrium of  $\mathcal{G}_M^{\text{EV}}$ .
- 4) For each game  $\mathcal{G}_M^{\text{EV}}$  the operator  $F_N$  is strongly monotone. Hence, Algorithm 2 converges to a variational Nash equilibrium of  $\mathcal{G}_M^{\text{EV}}$ .  $\square$

*Proof:* 1) We show that Assumption 1 holds. Indeed the sets  $\mathcal{X}^i$  in (27) are convex and compact, and  $\mathcal{Q}$  is non-empty by assumption. For each  $z$  fixed, the function  $J^i(x^i, z)$  is linear hence convex and continuously differentiable in  $x^i$ . We prove in the last statement that  $F_N$  is strongly monotone. This is equivalent to  $\nabla_x F_N(x) \succ 0$  by Lemma 2, which by definition of  $F_N(x)$  implies  $\nabla_{x^i}(\nabla_{x^i} J^i(x^i, \sigma(x))) \succ 0$ , which implies convexity of  $J^i(x^i, \sigma(x))$ . Finally,  $J^i(x^i, \sigma(x))$  is continuously differentiable because  $p_t$  is twice continuously differentiable. Having verified Assumption 1, Lemma 1 guarantees the existence of a Nash and of a Wardrop equilibrium.

The  $\varepsilon$ -Nash property is guaranteed by Proposition 2 upon verifying Assumption 2. This holds because  $\cup_{i=1}^M \mathcal{X}^i \subseteq [0, \tilde{x}^0]^n$  and because  $p$  is Lipschitz in  $[0, \tilde{x}^0]^n$  since  $p_t$  is assumed Lipschitz in  $[0, \tilde{x}^0]$  for all  $t$ . We conclude by noting that  $R = \tilde{x}^0 \sqrt{n}$ .

2) The fact that each  $p_t$  is strictly increasing in  $[0, \tilde{x}^0]$  implies that  $\nabla_z p(z) \succ 0$  in  $[0, \tilde{x}^0]^n$ , where  $p(z) := [p_1(\frac{d_1+z_1}{\kappa}), \dots, p_n(\frac{d_n+z_n}{\kappa})]^\top$ . In turn  $\nabla_z p(z) \succ 0$  guarantees strong monotonicity of  $p$  in  $[0, \tilde{x}^0]^n$  by Lemma 2. This, together with Assumptions 1 and 2 verified above, allows us to use the third result in Theorem 1.

3) Since  $\mathcal{X}$  is closed and convex, [11, Theorem 12.1.11] guarantees that the extragradient algorithm converges to a Wardrop equilibrium if  $F_W$  is monotone, which follows from the first statement of Lemma 3.

4) We have proven in the third statement that  $F_W$  is monotone. According to (10b), to show strong monotonicity of  $F_N$  it is sufficient to show that under condition (30) the term  $[\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M$  is strongly monotone for all  $x \in \mathcal{X}$ , which is equivalent to  $\nabla_x [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M \succ 0$  for all  $x \in \mathcal{X}$  by Lemma 2. We have

$$\nabla_x [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M = I_M \otimes \nabla_z p(z)|_{z=\sigma(x)} + \frac{1}{M} \mathbb{1}_M \otimes ([\text{diag}\{p''_t(\sigma_t) x_t^i\}_{t=1}^n]_{i=1}^M)^\top, \quad (31)$$

where  $\text{diag}\{p''_t(\sigma_t) x_t^i\}_{t=1}^n$  is the diagonal matrix whose entry in position  $(t, t)$  is  $p''_t(\sigma_t) x_t^i$ . The permutation matrix  $P = [[e_{t+(i-1)n}^\top]_{i=1}^M]_{t=1}^n$  permutes (31) into block-diagonal form

$$P \nabla_x [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M P^\top = \begin{bmatrix} p'_1(\sigma_1) I_M & & \\ & \ddots & \\ & & p'_n(\sigma_n) I_M \end{bmatrix} + \frac{1}{M} \begin{bmatrix} p''_1(\sigma_1) x_1 \mathbb{1}_M^\top & & \\ & \ddots & \\ & & p''_n(\sigma_n) x_n \mathbb{1}_M^\top \end{bmatrix} \quad (32)$$

where  $x_t = [x_t^i]_{i=1}^M$ . It suffices to show  $p'_t(\sigma_t) I_M + \frac{1}{M} p''_t(\sigma_t) x_t \mathbb{1}_M^\top \succ 0$  for all  $t$ . By Lemma 5 in Appendix,  $\lambda_{\min}(x_t \mathbb{1}_M^\top + \mathbb{1}_M x_t^\top) / 2 \geq -\frac{\tilde{x}^0 M}{8}$ , which ends the proof<sup>4</sup>.  $\blacksquare$

The average population strategy plays an important role in the EV application: indeed, [10, Theorem 6.1] shows in the same game setup that the average population strategy relative to a Nash equilibrium presents desirable properties for the grid operator. Nonetheless, if condition (30) is not satisfied, a Nash equilibrium cannot be achieved; it is instead possible to achieve a Wardrop equilibrium with the extragradient algorithm. The second statement of Corollary 1 then provides guarantees on the distance between the average population strategies at the Nash and at the Wardrop equilibrium.

#### Uniqueness of dual variables.

Corollary 1 shows that under condition (30) the operator  $F_N$  of  $\mathcal{G}_M^{\text{EV}}$  is strongly monotone, hence the game  $\mathcal{G}_M^{\text{EV}}$  admits a unique variational Nash equilibrium by Lemma 1. We study here the uniqueness of the associated dual variables  $\bar{\lambda}_N$  introduced in Proposition 3. Guaranteeing unique dual variables might be important to convince the vehicle owners

<sup>4</sup>The work [42] studies an aggregative game and in [42, Lemma 3] it exploits expression (32) to give conditions for  $\nabla_x F_N(x)$  to be a  $P$ -matrix, which in turn guarantees uniqueness of the Nash equilibrium in absence of coupling constraints. It is interesting to note that uniqueness in [42] holds assuming  $p'_t > 0, p''_t > 0$ , whereas for us it suffices  $p'_t > 0, p''_t < 0$ .



to participate in the proposed scheme, as the dual variables represent the penalty price associated to the coupling constraint. Define  $R^{\text{tight}} \subseteq \{1, \dots, n\}$  as the set of instants in which the coupling constraint  $\mathcal{C}$  is active. We provide a sufficient condition for uniqueness of the dual variables which relies on a slight modification of the linear-independence constraint qualification [43].

**Proposition 4.** *Assume that condition (30) holds and consider the unique variational Nash equilibrium  $\bar{x}_N$  of  $\mathcal{G}_M^{\text{EV}}$ . If there exists a vehicle  $i$  such that*

- $\bar{x}_{N,t}^i \notin \{0, \tilde{x}_t^i\}$  for all  $t \in R^{\text{tight}}$  and
- $\bar{x}_{N,t'}^i \notin \{0, \tilde{x}_{t'}^i\}$  for some  $t' \notin R^{\text{tight}}$ ,

*then the dual variables  $\bar{\lambda}_N$  associated to the coupling constraint (28) are unique.*  $\square$

The proof is reported in the Appendix. We note that the sufficient condition of Proposition 4 is to be verified a-posteriori; in other words, it depends on the primal solution  $\bar{x}_N$ . In the numerical analysis presented in the following such sufficient condition always holds. Uniqueness of the dual variables associated to the coupling constraint of an aggregative game has been studied also in [42, Theorem 4], where the conditions in the bullets of Proposition 4 are not required but  $p$  is restricted to be affine.

### B. Numerical analysis

The numerical study is conducted on a heterogeneous population of agents. We set the price function to  $p_t(z_t) = 0.15\sqrt{(d_t + \sigma_t(x))/\kappa_t}$  as in [10, eq.(25)], and  $n = 24$ . The agents differ in  $\theta^i$ , randomly chosen according to  $\mathcal{U}[0.5, 1.5]$ ; they also differ in  $\tilde{x}_t^i$ , which is chosen such that the charge is allowed in a connected interval, with left and right endpoints uniformly randomly chosen: within the interval,  $\tilde{x}_t^i$  is constant and randomly chosen for each agent, according to  $\mathcal{U}[1, 5]$ ; outside this interval,  $\tilde{x}_t^i = 0$ . The demand  $d_t$  is taken as the typical (non-EV) base demand over a summer day in the United States [10, Figure 1];  $\kappa_t = 12$  kW for all  $t$ , and the upper bound  $K_t = 0.55$  kW is chosen such that the coupling constraint (28) is active in the middle of the night. Note that with these choices all the assumptions of Corollary 1 are met. In particular, for the given choice of  $p$  condition (30) holds because  $p_t''(z) < 0$  for all  $z$  and all  $t$ . Figure 1 presents the aggregate consumption at the Nash equilibrium found by Algorithm 2, with stopping criterion  $\|(x_{(k+1)}, \lambda_{(k+1)}) - (x_{(k)}, \lambda_{(k)})\|_\infty \leq 10^{-4}$ .

Note that without the coupling constraint the quantity  $\bar{\sigma} + d$  would be constant overnight, as shown in [10]. Figure 2 illustrates the bound  $\|\sigma(\bar{x}_N) - \sigma(\bar{x}_W)\| \leq \tilde{x}^0 \sqrt{\frac{2nL_p}{\alpha M}}$  of the second statement of Corollary 1. The Wardrop equilibrium is computed with the extragradient algorithm with stopping criterion  $\|(x_{(k+1)}, \lambda_{(k+1)}) - (x_{(k)}, \lambda_{(k)})\|_\infty \leq 10^{-4}$ . The  $\varepsilon$ -Nash property of the Wardrop equilibrium in Proposition 2 can also be illustrated; a plot is omitted here for reasons of space.

The framework introduced above can also be used to enforce local coupling constraints, i.e. constraints on a subset of all the vehicles. These can for instance be used to model capacity limits

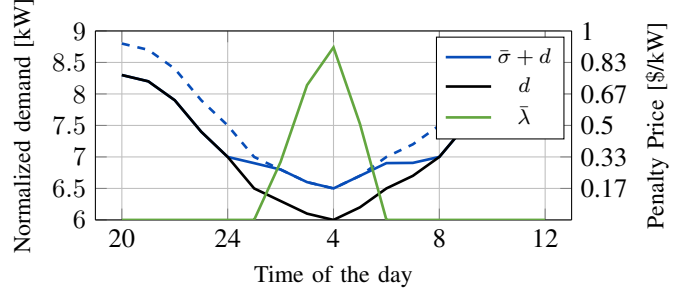


Fig. 1: Aggregate EV demand  $\sigma(\bar{x}_N)$  and dual variables  $\bar{\lambda}_N$  for  $M = 100$ , subject to  $\sigma(x) \leq 0.55$  kW. The region below the dashed line corresponds to  $\sigma(x) + d \leq 0.55$  kW +  $d$ .

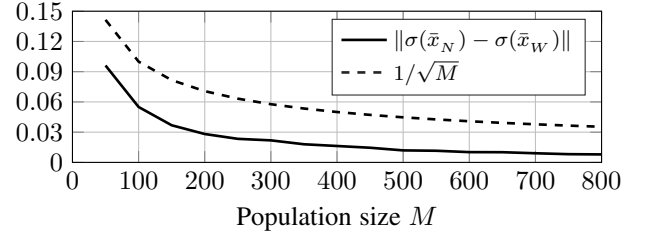


Fig. 2: Distance between the aggregates  $\sigma(\bar{x}_N)$  and  $\sigma(\bar{x}_W)$ .

for local substations. We refer the reader to [6, Section VI] for a more detailed analysis.

### Quadratic cost function

Different works in the EV literature [14], [44] use the quadratic cost (25), with  $Q \succ 0$  and  $C \succ 0$ , diagonal. Existence of a Nash and of a Wardrop equilibrium is guaranteed by Lemma 1, while Proposition 2 gives the  $\varepsilon$ -Nash property. Further, Lemma 4 shows that the resulting operators  $F_N$  and  $F_W$  are strongly monotone with monotonicity constant independent from  $M$ . Theorem 1 ensures then that  $\|\bar{x}_N - \bar{x}_W\| \leq (L_p R)/(\alpha \sqrt{M})$ . A Nash equilibrium can be found using Algorithm 1, while a Wardrop equilibrium can be achieved using both Algorithm 1 or 2. Figure 3 presents a comparison between the two algorithms in terms of iteration count, where  $Q = 0.1I_n$ ,  $C = I_n$ ,  $c^i = d$  for all  $i$ . Figure 3 (top) represents the number of strategy updates required to converge, i.e. the number of times (21) or (23b) is used. Figure 3 (bottom) depicts the number of dual variables updates, i.e. the number of times (22) or (23c) is used. For both algorithms the number of iterations does not seem to increase with the population size. Algorithm 2 requires fewer primal iterations, while Algorithm 1 needs much fewer dual iterations.

## VII. ROUTE CHOICE IN A ROAD NETWORK

As second application we study a population of drivers interacting in a road network. Our model differs from [45] in the cost function (35), where we introduce a term penalizing the deviation from a preferred route. We assume that the travel time on each road depends only on the traffic on that road, whereas [26] considers also upstream and downstream influence. While most traffic literature focuses solely on the Wardrop

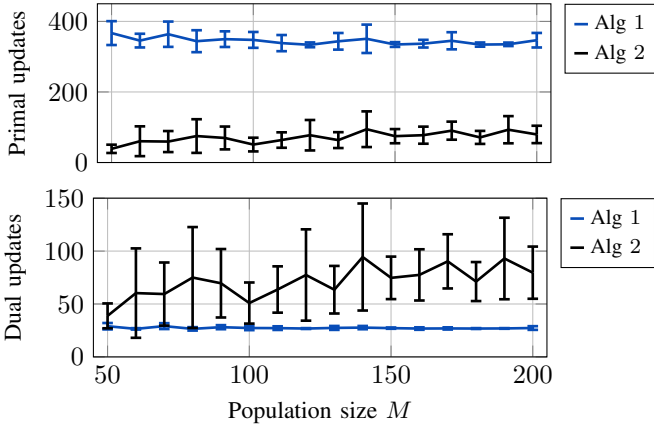


Fig. 3: Primal (top) and dual (bottom) updates required to converge; mean and standard deviation for 10 repetitions. As each step of Algorithm 2 performs one primal and one dual update, the two black lines (top and bottom) coincide.

equilibrium [45], [26], we also study the Nash equilibrium and illustrate the distance between the two.

We consider a strongly-connected directed graph  $(\mathcal{V}, \mathcal{E})$  with vertex set  $\mathcal{V} = \{1, \dots, V\}$ , representing geographical locations, and directed edge set  $\mathcal{E} = \{1, \dots, E\} \subseteq \mathcal{V} \times \mathcal{V}$ , representing roads connecting the locations. Each agent  $i \in \{1, \dots, M\}$  represents a driver who wants to drive from his origin  $o^i \in \mathcal{V}$  to his destination  $d^i \in \mathcal{V}$ .

### Constraints

Let us introduce the vector  $x^i \in [0, 1]^E$  to describe the strategy (route choice) of agent  $i$ , with  $[x^i]_e$  representing the probability that agent  $i$  transits on edge  $e$  [46]. To guarantee that agent  $i$  leaves his origin and reaches his destination with probability 1, the strategy  $x^i$  has to satisfy

$$\sum_{e \in \text{in}(v)} [x^i]_e - \sum_{e \in \text{out}(v)} [x^i]_e = \begin{cases} -1 & \text{if } v = o^i \\ 1 & \text{if } v = d^i \\ 0 & \text{otherwise,} \end{cases} \quad \forall v \in \mathcal{V},$$

where  $\text{in}(v)$  and  $\text{out}(v)$  represent the set of in-edges and the set of out-edges of node  $v$ . We denote the graph incidence matrix by  $B \in \mathbb{R}^{V \times E}$ , so that  $[B]_{ve} = 1$  if edge  $e$  points to vertex  $v$ ,  $[B]_{ve} = -1$  if edge  $e$  exits vertex  $v$  and  $[B]_{ve} = 0$  otherwise. The individual constraint set of player  $i$  is then

$$\mathcal{X}^i := \{x \in [0, 1]^E : Bx = b^i\}, \quad (33)$$

where  $b^i \in \mathbb{R}^V$  is such that  $[b^i]_v = -1$  if  $v = o^i$ ,  $[b^i]_v = 1$  if  $v = d^i$  and  $[b^i]_v = 0$  otherwise. We introduce the coupling constraint

$$x \in \mathcal{C} := \{x \in \mathbb{R}^{ME} \mid \frac{1}{M} \sum_{i=1}^M x_e^i \leq K_e, \forall e = 1, \dots, E\}, \quad (34)$$

expressing the fact that the number of vehicles on edge  $e$  cannot exceed  $MK_e$ . Such constraint can be imposed by authorities to decrease the congestion in a specific road or neighborhood, with the goal of reducing noise or pollution.

### Cost function

We assume that each driver  $i \in \{1, \dots, M\}$  wants to minimize his travel time and, at the same time, does not want to deviate too much from a preferred route  $\tilde{x}^i \in \mathcal{X}^i$ . We model this objective with the following cost function

$$J^i(x^i, \sigma(x)) = \frac{\gamma^i}{2} \|x^i - \tilde{x}^i\|^2 + \sum_{e=1}^E t_e(\sigma_e(x_e)) x_e^i, \quad (35)$$

with  $\gamma^i \geq 0$  a weighting factor,  $x_e := [x_e^1, \dots, x_e^M]^\top$ ,  $\sigma_e(x_e) = \frac{1}{M} \sum_{i=1}^M x_e^i$  and  $t_e(\sigma_e(x_e))$  the travel time on edge  $e$ .

### Travel time

This subsection is devoted to the derivation of the analytical expression of the travel time  $t_e(\sigma_e(x_e))$ . The reader not interested in the technical details of the derivation can jump to the expression of  $t_e(\sigma_e(x_e))$  in (38), which is illustrated in Figure 4. We introduce the quantity  $D_e(x_e) = \sum_{i=1}^M x_e^i$  to describe the total demand on edge  $e$ . We consider a rush-hour interval  $[0, h]$  and we assume that the instantaneous demand equals  $D_e(x_e)/h$  at any time  $t \in [0, h]$  and zero for  $t > h$ . We assume that edge  $e$  can support a maximum flow  $F_e$  (vehicles per unit of time) and features a free-flow travel time  $t_{e,\text{free}}$ . As we are interested in comparing populations of different sizes, we further assume that the peak hour duration  $h$  is independent from the population size  $M$  and that the road maximum capacity flow  $F_e$  scales linearly with the population size, i.e.  $F_e(M) = f_e \cdot M$ , with  $f_e$  constant in  $M$ . The consideration underpinning this last assumption is that the road infrastructure scales with the number of vehicles to accommodate the increasing demand, similarly as what assumed in [10] for the energy infrastructure.

If  $D_e(x_e)/h \leq F_e$  then every car has instantaneous access to edge  $e$  and no queue accumulates, hence the travel time equals  $t_{e,\text{free}}$ . We focus in the rest of this paragraph on the case  $D_e(x_e)/h > F_e$ . An increasing queue forms in the interval  $[0, h]$  and decreases at rate  $F_e$  for  $t > h$ . The number of vehicles  $q_e(t)$  queuing on edge  $e$  at time  $t$  obeys then the dynamics

$$\dot{q}_e(t) = \begin{cases} \frac{D_e(x_e)}{h} \cdot \mathbf{1}_{[0,h]}(t) - F_e & \text{if } q_e(t) \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad q_e(0) = 0, \quad (36)$$

where  $\mathbf{1}_{[0,h]}$  is the indicator function of  $[0, h]$ . The solution  $q_e(t)$  to (36) is hence

$$q_e(t) = \begin{cases} \left( \frac{D_e(x_e) - F_e h}{h} \right) t & \text{if } 0 \leq t \leq h \\ D_e(x_e) - F_e t & \text{if } h \leq t \leq D_e(x_e)/F_e \\ 0 & \text{if } t \geq D_e(x_e)/F_e. \end{cases} \quad (37)$$

As a consequence, the total queuing time at edge  $e$  (i.e., the queuing times summed over all vehicles) is the integral of  $q_e(t)$ , which equals  $D_e(x_e)(D_e(x_e) - F_e h)/(2F_e)$ ; the queuing time is then  $(D_e(x_e) - F_e h)/(2F_e)$ .

Since  $\sigma_e(x_e) = \frac{1}{M} \sum_{i=1}^M x_e^i = \frac{1}{M} D_e(x_e)$ , the travel time is

$$t_e^{\text{PWA}}(\sigma_e(x_e)) = \begin{cases} t_{e,\text{free}} & \text{if } \sigma_e(x_e) \leq f_e h \\ t_{e,\text{free}} + \frac{\sigma_e(x_e) - f_e h}{2f_e} & \text{otherwise,} \end{cases}$$

and is reported in Figure 4. Note that  $t_e^{\text{PWA}}$  is a continuous and piece-wise affine function of  $\sigma_e(x_e)$ , but it is not continuously

differentiable, hence Assumption 1 would not hold. Therefore, we define  $t_e$  appearing in (35) as the smoothed version of  $t_e^{\text{PWA}}$

$$t_e(\sigma_e(x_e)) = \begin{cases} t_{e,\text{free}} & \text{if } \sigma_e(x_e) \leq f_e h - \Delta_e \\ t_{e,\text{free}} + \frac{\sigma_e(x_e) - f_e h}{2f_e} & \text{if } \sigma_e(x_e) \geq f_e h + \Delta_e \\ a\sigma_e(x_e)^2 + b\sigma_e(x_e) + c & \text{otherwise,} \end{cases} \quad (38)$$

where the values of  $\Delta_e$ ,  $a$ ,  $b$ ,  $c$  are such that  $t_e$  is continuously differentiable<sup>5</sup>, as illustrated in Figure 4.

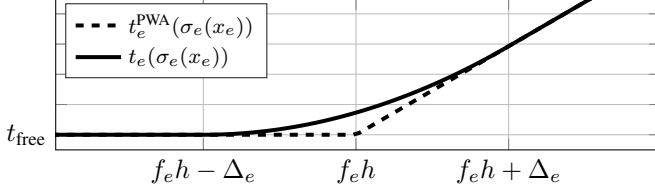


Fig. 4: Piece-wise affine travel time  $t_e^{\text{PWA}}(\sigma_e(x_e))$  and its smooth approximation  $t_e(\sigma_e(x_e))$  as functions of  $\sigma_e(x_e)$ .

We note that the function  $t_e(\sigma_e(x_e))$  is used within a stationary traffic model but includes the average queuing time which is based on the dynamic function (37). A thorough analysis of a dynamic traffic model is subject of future work.

Finally, we remark that a travel time with similar monotonicity properties can be derived from the piecewise affine fundamental diagram of traffic [47, Figure 7], but  $t_e(\sigma_e(x_e))$  would present a vertical asymptote which is absent here.

#### A. Theoretical guarantees

We define the route-choice game  $\mathcal{G}_M^{\text{RC}}$  as in (3), with  $\mathcal{X}^i$  as in (33),  $\mathcal{C}$  as in (34) and  $J^i(x^i, \sigma(x))$  as in (35), (38). In the following we apply the main results of Sections III, IV, V to the route choice game.

**Corollary 2.** Consider the sequence of games  $(\mathcal{G}_M^{\text{RC}})_{M=1}^\infty$ . Assume that for each game  $\mathcal{G}_M^{\text{RC}}$  the set  $\mathcal{Q} = \mathcal{C} \cap \mathcal{X}$  is non-empty, that  $h > 0$  and  $t_{e,\text{free}}, f_e > 0$  for each  $e \in \mathcal{E}$ . Moreover, assume that there exists  $\hat{\gamma} > 0$  such that  $\gamma^i \geq \hat{\gamma}$  for all  $i \in \{1, \dots, M\}$ , for all  $M$ . Then:

- 1) The operator  $F_W$  is strongly monotone, hence each game  $\mathcal{G}_M^{\text{RC}}$  admits a unique variational Wardrop equilibrium. For every  $M$  satisfying

$$M > \max_{e \in \mathcal{E}} \frac{1}{32f_e \Delta_e \hat{\gamma}} \quad (39)$$

the operator  $F_N$  is strongly monotone, hence each game  $\mathcal{G}_M^{\text{RC}}$  admits a unique variational Nash equilibrium. Every Wardrop equilibrium is an  $\varepsilon$ -Nash equilibrium with  $\varepsilon = \frac{E}{M f_{\min}}$ , where  $f_{\min} = \min_{e \in \mathcal{E}} f_e$ .

- 2) For any variational Nash equilibrium  $\bar{x}_N$  of  $\mathcal{G}_M^{\text{RC}}$ , the unique variational Wardrop equilibrium  $\bar{x}_W$  of  $\mathcal{G}_M^{\text{RC}}$  satisfies

$$\|\bar{x}_N - \bar{x}_W\| \leq \frac{\sqrt{E}}{2f_{\min} \hat{\gamma} \sqrt{M}}.$$

<sup>5</sup>The values are  $\Delta_e = 0.5(\sqrt{(f_e h)^2 + 4f_e h} - f_e h)$ ,  $a = 1/(8f_e \Delta_e)$ ,  $b = 1/(4f_e) - h/(4\Delta_e)$ ,  $c = t_{e,\text{free}} + (f_e h)^2/(8f_e \Delta_e) - h/4 - (\Delta_e)/(8f_e)$ .

- 3) For any  $M$ , Algorithm 2 with operator  $F_W$  converges to a variational Wardrop equilibrium of  $\mathcal{G}_M^{\text{RC}}$ . For  $M$  satisfying (39), Algorithm 2 with operator  $F_N$  converges to a variational Nash equilibrium of  $\mathcal{G}_M^{\text{RC}}$ .  $\square$

*Proof:* 1) Assumption 1 and the consequent existence of a variational Nash and of a variational Wardrop equilibrium for any  $M$  can be shown as in Corollary 1. The operator  $F_W$  for the cost (35) reads

$$F_W(x) = [\gamma^i(x^i - \hat{x}^i) + t(\sigma(x))]_{i=1}^M.$$

where  $t(\sigma(x)) := [t_e(\sigma_e(x_e))]_{e=1}^E$ . Since  $t_e(\sigma_e(x_e))$  in (38) is a monotone function of  $\sigma_e(x_e)$ , the operator  $t(\sigma(x))$  is monotone. Then  $F_W$  is strongly monotone with constant  $\hat{\gamma}$  because it is the sum of a monotone and a strongly monotone operator with constant  $\hat{\gamma}$ . As a consequence, each  $\mathcal{G}_M^{\text{RC}}$  admits a unique variational Wardrop equilibrium.

To prove strong monotonicity of  $F_N$  we use the result of Lemma 2<sup>6</sup>. We first note that each  $t_e$  only depends on the corresponding  $\sigma_e$ , hence  $\nabla_x F_N(x)$  can be permuted into diagonal form similarly to what done in (32). It then suffices to show  $\hat{\gamma}I_M + \frac{1}{M}t'_e(\sigma_e)I_M + \frac{1}{M^2}t''_e(\sigma_e)x_e \mathbb{1}_M^\top \succ 0$  for all  $\sigma_e$  and for all  $e$ . This matrix is indeed positive definite if  $\sigma_e(x_e) \notin [f_e h - \Delta_e, f_e h + \Delta_e]$ , because then  $t'_e(\sigma_e) \geq 0$  and  $t''_e(\sigma_e) = 0$  by (38). For  $\sigma_e(x_e) \in [f_e h - \Delta_e, f_e h + \Delta_e]$  it suffices to show  $\hat{\gamma}I_M + \frac{1}{M^2}t''_e(\sigma_e)x_e \mathbb{1}_M^\top \succ 0$ , because  $t'_e(\sigma_e) \geq 0$  and  $t''_e(\sigma_e) = \frac{1}{4f_e \Delta_e}$ . By Lemma 5 in the Appendix,  $\lambda_{\min}(x_e \mathbb{1}_M^\top + \mathbb{1}_M x_e^\top)/2 \geq -\frac{M}{8}$ , which proves strong monotonicity of  $F_N$  under (39). Consequently, if  $M$  satisfies (39) then  $\mathcal{G}_M^{\text{RC}}$  admits a unique variational Nash equilibrium. Finally, we verify Assumption 2 in order to use Proposition 2. We have  $\mathcal{X}^0 = [0, 1]^E$  and  $t$  is continuously differentiable and hence Lipschitz in  $\mathcal{X}^0$ , with Lipschitz constant  $L_p = 1/(2f_{\min})$ . Moreover,  $R := \max_{y \in \mathcal{X}^0} \{\|y\|\} = \sqrt{E}$ .

2) Since all the assumptions of Theorem 1 have just been verified, it is a direct consequence of its second statement.

3) Since all the assumptions of Theorem 3 have just been verified, it is a direct consequence of its statement.  $\blacksquare$

#### B. Numerical analysis

For the numerical analysis we use the data set of the city of Oldenburg [27], whose graph features 175 nodes and 213 undirected edges<sup>7</sup> and is reported in Figure 5. For each agent  $i$  the origin  $o^i$  and the destination  $d^i$  are chosen uniformly at random. Regarding the cost (35),  $t_{e,\text{free}}$  is computed as the ratio between the road length, which is provided in the data set, and the free-flow speed. Based on the road topology, we divide the roads into main roads, where the free-flow speed is 50 km/h, and secondary roads, where the free-flow speed is 30 km/h. Moreover, we assume a peak hour duration  $h$  of 2 hours, and

<sup>6</sup>Lemma 2 requires  $F_N$  to be continuously differentiable, which is not the case here. The more general result [48, Proposition 2.1] extends the statement of Lemma 2 to operators which are not continuously differentiable. It then suffices to show  $\nabla_x F_N(x) \succ 0$  for  $\sigma(x)$  in each of the three intervals defined by (38), because in each of them  $F_N$  is continuously differentiable.

<sup>7</sup>The graph in the original data set features 6105 vertexes and 7035 undirected edges. We reduce it by excluding all the nodes that are outside the rectangle  $[3619, 4081] \times [3542, 4158]$  and all the edges that do not connect two nodes in the rectangle. The resulting graph is strongly connected.

for all  $e \in \mathcal{E}$ , we set  $f_e = 4 \cdot 10^{-3}$  vehicles per second, which corresponds to 1 vehicle every 4 seconds for a population of  $M = 60$  vehicles. Finally, the parameter  $\gamma^i$  is picked uniformly at random in  $[0.5, 3.5]$  and  $\tilde{x}^i$  is such that  $\tilde{x}_e^i = 1$  if  $e$  belongs to the shortest path from  $o^i$  to  $d^i$ , while  $\tilde{x}_e^i = 0$  otherwise. The shortest path is computed based on  $\{t_{e,\text{free}}\}_{e=1}^E$ . Note that with the above values the bound (39) becomes  $M > 16.14$ , which is satisfied also for small-size populations.

We compute the Wardrop equilibrium with Algorithm 2 relatively to a population of  $M = 60$  drivers without coupling constraint, i.e. with  $K_e = 1$  for all  $e \in \mathcal{E}$ . We report in Figure 5 the corresponding queuing time  $t_e(\sigma_e(x_e)) - t_{e,\text{free}}$  as by (38).

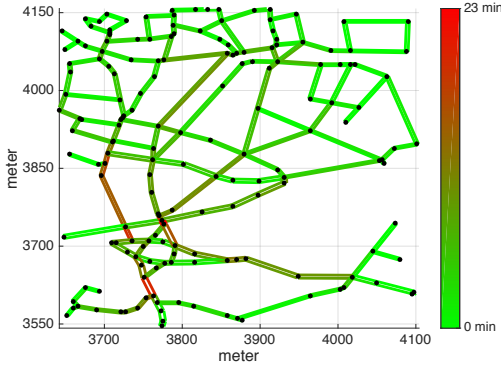


Fig. 5: The queuing time reported in green-red color scale. Note that this pattern changes if one modifies the pairs origin-destination.

We illustrate in Figure 6 the change in the queuing time of an entire neighborhood when introducing a coupling constraint that upper bounds the total number of cars on a single edge, relatively to a Wardrop equilibrium with  $M = 60$ .

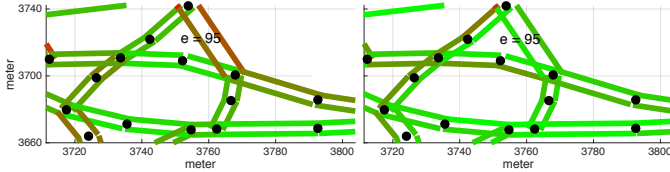


Fig. 6: On the left, the queuing time in a neighborhood without any coupling constraints; 10% of the population transits on edge 95, and the queuing time is 7.28 minutes. On the right, the queuing time in presence of a coupling constraint allowing at most 3% of the entire population on edge 95; the queuing time is reduced to 1.42 minutes, but it visibly increases on the edges of the alternative route.

Finally, we illustrate the second statement of Corollary 2 by reporting in Figure 7 the distance between the unique variational Wardrop equilibrium and the variational Nash equilibrium found by Algorithm 2. The  $\varepsilon$ -Nash property of the Wardrop equilibrium in Proposition 2 can also be illustrated, but a plot is omitted here for reasons of space.

## VIII. CONCLUSIONS

The paper considered aggregative games and established novel results on the Euclidean distance between Nash and Wardrop equilibrium; moreover, it proposed two decentralized

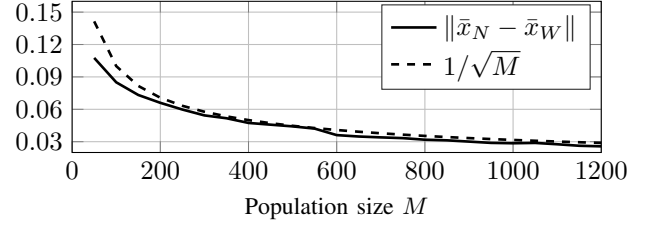


Fig. 7: Distance between Nash and Wardrop variational equilibria.

algorithms to achieve the two equilibria in presence of coupling constraints and investigated two relevant applications. As future research direction, it would be interesting to design distributed algorithms which achieve an equilibrium by means of local communications. Moreover, by exploiting the VI reformulation one could establish results on the proximity between Nash equilibrium and social optimum. Finally, vehicle dynamics could be included in the road network case study of Section VII, thus making the model more realistic.

## APPENDIX: PROOFS

### Proof of Lemma 3

1) Let us first show that  $F_W$  is monotone. Since  $v^i$  is convex, then  $\nabla_{x^i} v^i(x^i)$  is monotone in  $x^i$  by [36, Section 4.2.2]. Hence  $[\nabla_{x^i} v^i(x^i)]_{i=1}^M$  is monotone. Moreover, for any  $x_1, x_2$

$$\begin{aligned} & ([p(\sigma(x_1))]_{i=1}^M - [p(\sigma(x_2))]_{i=1}^M)^\top (x_1 - x_2) \\ &= M(p(\sigma(x_1)) - p(\sigma(x_2)))^\top (\sigma(x_1) - \sigma(x_2)) \geq 0, \end{aligned} \quad (40)$$

where the last inequality follows from the fact that  $p$  is monotone. By (10a) and the fact that the sum of two monotone operators is monotone, one can conclude that  $F_W$  is monotone. To show that  $F_N$  is strongly monotone, we write the affine expression of  $p$  as  $p(x) = Cx + c$ , where there exists  $\alpha > 0$  such that  $C \succ \alpha I_n$  by Lemma 2. Then the term  $\frac{1}{M} [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M$  in (10b) equals  $\frac{1}{M} (I_M \otimes C^\top)x$ . Since  $\nabla_x (\frac{1}{M} (I_M \otimes C^\top)x) \succ \frac{\alpha}{M} I_{Mn}$ , then  $\frac{1}{M} [\nabla_z p(z)|_{z=\sigma(x)} x^i]_{i=1}^M$  is strongly monotone by Lemma 2. Having already shown that  $F_W$  is monotone, the proof is concluded upon noting that the sum of a monotone operator and a strongly monotone operator is strongly monotone.

2) Strong convexity of  $v^i$  is equivalent to strong monotonicity of  $\nabla_{x^i} v^i(x^i)$  in  $x^i$  [36, Section 4.2.2]. Then  $[\nabla_{x^i} v^i(x^i)]_{i=1}^M$  is strongly monotone. Monotonicity of  $[p(\sigma(x))]_{i=1}^M$  in (10a) can be shown as in (40). ■

### Proof of Proposition 3

Under Assumptions 1 and 3 the set  $\mathcal{Q}$ , and consequently the sets  $\{\mathcal{X}^i\}_{i=1}^M$ ,  $\mathcal{X}$  and  $\mathcal{Y}$ , are convex and satisfy Slater's constraint qualification. The VI( $\mathcal{Q}, F$ ) is therefore equivalent to its KKT system [11, Proposition 1.3.4]. Moreover, since  $\mathcal{X}^i$  satisfies Slater's constraint qualification, the optimization problem of agent  $i$  in the game (18) is equivalent to its KKT system, for each  $i$ . Finally, by [11, Proposition 1.3.4], the VI( $\mathcal{Y}, T$ ) is equivalent to its KKT system. We do not report the three KKT systems here, but it can be seen by direct inspection that they are equivalent [36, Section 4.3.2]. ■

### Proof of Theorem 2

We split the proof of the theorem into two parts. First we show convergence of the inner loop and then of the outer loop.

*Inner loop.* Using the same approach of [14, Theorem 3 and Corollary 1], it is possible to show that under Assumption 4 for any  $\lambda_{(k)} \in \mathbb{R}_{\geq 0}^m$  the sequences of  $z_{(h)}$  and of  $\tilde{x}(h)$  converge respectively to  $\bar{z}$  and to  $\bar{x}$  such that  $\bar{z} = \frac{1}{M} \sum_{i=1}^M x_{\text{or}}^i(\bar{z}, \lambda_{(k)}) =: \frac{1}{M} \sum_{i=1}^M \bar{x}^i = \sigma(\bar{x})$ . In [14, Theorem 1] it is shown that the set  $\{\bar{x}^i\}_{i=1}^M$  is an  $\varepsilon$ -Nash equilibrium for the game  $\mathcal{G}(\lambda_{(k)})$ , with  $\varepsilon = \mathcal{O}(\frac{1}{M})$ . In the following, we show that  $\{\bar{x}^i\}_{i=1}^M$  is actually a Wardrop equilibrium of  $\mathcal{G}(\lambda_{(k)})$ <sup>8</sup>. Indeed, for each agent  $i$ , by the definition of optimal response in (20), one has

$$J^i(\bar{x}^i, \bar{z}) + \lambda_{(k)}^\top A_{(:,i)} \bar{x}^i \leq J^i(x^i, \bar{z}) + \lambda_{(k)}^\top A_{(:,i)} x^i, \forall x^i \in \mathcal{X}^i.$$

Using the fact that  $\bar{z} = \sigma(\bar{x})$ , we get

$$J^i(\bar{x}^i, \sigma(\bar{x})) + \lambda_{(k)}^\top A_{(:,i)} \bar{x}^i \leq J^i(x^i, \sigma(\bar{x})) + \lambda_{(k)}^\top A_{(:,i)} x^i,$$

for all  $x^i \in \mathcal{X}^i$  and for all  $i \in \{1, \dots, M\}$ . Thus  $\{\bar{x}^i\}_{i=1}^M$  is a Wardrop equilibrium of  $\mathcal{G}(\lambda_{(k)})$  by Definition 2.

*Outer loop.* We follow the steps of the proof of [38, Proposition 8]. For each  $\lambda \in \mathbb{R}_{\geq 0}^m$  define  $F_W(x; \lambda) := F_W(x) + A^\top \lambda$ . Such operator is strongly monotone in  $x$  on  $\mathcal{Q}$  with the same constant  $\alpha$  as  $F_W(x)$ . It follows by Lemma 1, that  $\mathcal{G}(\lambda)$  has a unique variational Wardrop equilibrium which we denote by  $\bar{x}_W(\lambda)$ . Note that the outer loop update can be written as

$$\lambda_{(k+1)} = \Pi_{\mathbb{R}_{\geq 0}^m} [\lambda_{(k)} - \tau(b - A\bar{x}_W(\lambda_{(k)}))],$$

which is a step of the projection algorithm [11, Algorithm 12.1.4] applied to  $\text{VI}(\mathbb{R}_{\geq 0}^m, \Phi)$ , with  $\Phi(\lambda) := b - A\bar{x}_W(\lambda)$ . To conclude, it suffices to show that  $\lambda_{(k)}$  converges to a solution  $\bar{\lambda}$  of such VI, because by [11, Proposition 1.1.3],  $\bar{\lambda}$  solves  $\text{VI}(\mathbb{R}_{\geq 0}^m, \Phi)$  if and only if  $0 \leq \bar{\lambda} \perp (b - A\bar{x}_W(\bar{\lambda})) \geq 0$ . Having already proved convergence of the inner loop, the conclusion then follows from the second statement of Proposition 3.

To show that the sequence  $\lambda_{(k)}$  converges to a solution of the  $\text{VI}(\mathbb{R}_{\geq 0}^m, \Phi)$ , we prove that the mapping  $\Phi$  is co-coercive<sup>9</sup> with co-coercitivity constant  $c_\Phi = \alpha/\|A\|^2$  and apply [11, Theorem 12.1.8] to conclude the proof. Note that [11, Theorem 12.1.8] requires  $\text{VI}(\mathbb{R}_{\geq 0}^m, \Phi)$  to have at least a solution; this is guaranteed by the equivalence between 1) and 2) in Proposition 3 upon noting that a solution of  $\text{VI}(\mathcal{Q}, F)$  exists by Lemma 1.

To show co-coercitivity of  $\Phi$ , consider  $\lambda_1, \lambda_2 \in \mathbb{R}_{\geq 0}^m$  and the corresponding unique solutions  $x_1 := \bar{x}_W(\lambda_1)$  of  $\text{VI}(\mathcal{X}, F_W + A^\top \lambda_1)$  and  $x_2 := \bar{x}_W(\lambda_2)$  of  $\text{VI}(\mathcal{X}, F_W + A^\top \lambda_2)$ . By definition

$$(x_2 - x_1)^\top (F_W(x_1) + A^\top \lambda_1) \geq 0, \quad (41a)$$

$$(x_1 - x_2)^\top (F_W(x_2) + A^\top \lambda_2) \geq 0. \quad (41b)$$

Adding (41a) and (41b) we obtain  $(x_2 - x_1)^\top (F_W(x_1) - F_W(x_2) + A^\top (\lambda_1 - \lambda_2)) \geq 0$ , i.e.,  $(x_2 - x_1)^\top A^\top (\lambda_1 - \lambda_2) \geq (x_2 - x_1)^\top (F_W(x_2) - F_W(x_1))$ . Since  $F_W$  is strongly monotone, it follows from the last inequality that

$$(Ax_2 - Ax_1)^\top (\lambda_1 - \lambda_2) \geq \alpha \|x_2 - x_1\|^2. \quad (42)$$

<sup>8</sup>This is consistent with [14, Theorem 1] thanks to Proposition 2.

<sup>9</sup>The operator  $\Phi : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is co-coercive with constant  $\eta > 0$  if  $(\Phi(\lambda_1) - \Phi(\lambda_2))^\top (\lambda_1 - \lambda_2) \geq \eta \|\Phi(\lambda_1) - \Phi(\lambda_2)\|^2$ , for all  $\lambda_1, \lambda_2 \in \mathbb{R}^m$ .

Moreover, since by definition of induced matrix norm  $\|A(x_2 - x_1)\| \leq \|A\| \|x_2 - x_1\|$ , then

$$\|x_2 - x_1\|^2 \geq \frac{\|A(x_2 - x_1)\|^2}{\|A\|^2}. \quad (43)$$

Combining (42), (43), and adding and subtracting  $b$ , we obtain

$$(b - Ax_2 - (b - Ax_1))^\top (\lambda_2 - \lambda_1) \geq \frac{\alpha}{\|A\|^2} \|b - Ax_2 - (b - Ax_1)\|^2,$$

hence  $\Phi$  is co-coercive in  $\lambda$  with constant  $c_\Phi = \alpha/\|A\|^2$ . ■

### Proof of Theorem 3

We give the proof for a strongly monotone operator  $F$ , which is to be interpreted as  $F_N$  in the first statement and  $F_W$  in the second statement. We divide the proof into two parts: (i) we prove that Algorithm 2 is a particular case of a class of algorithms known as asymmetric projection algorithms (APA) [11, Algorithm 12.5.1] applied to  $\text{VI}(\mathcal{Y}, T)$ ; (ii) we prove that our algorithm satisfies a convergence condition for APA.

(i) The APA are parametrized by the choice of a matrix  $D \succ 0$ . For a fixed  $D$  a step of the APA for  $\text{VI}(\mathcal{Y}, T)$  is

$$y_{(k+1)} = \text{solution of } \text{VI}(\mathcal{Y}, T_D^k), \quad (44)$$

where  $y_{(k)}$  is the state at iteration  $k$  and  $T_D^k(y) := T(y_{(k)}) + D(y - y_{(k)})$ . In other words every step of the APA requires the solution of a different variational inequality that depends on the operator  $T$ , on a fixed matrix  $D$  and on the previous strategies' vector  $y_{(k)}$ . We choose

$$D := \begin{bmatrix} \frac{1}{\tau} I_{Mn} & 0 \\ -2A & \frac{1}{\tau} I_m \end{bmatrix}, \quad (45)$$

which by using the Schur complement condition can be shown to positive definite under (24a). It is shown in [11, Section 12.5.1] that with the choice (45) the update (44) coincides with the steps (23).

(ii) As illustrated in the previous point, Algorithm 2 is the specific APA associated with the choice of  $D$  given in (45). According to [11, Proposition 12.5.2], this algorithm converges if the mapping  $G(y) = D_s^{-1/2} T(D_s^{-1/2} y) - D_s^{-1/2} (D - D_s) D_s^{-1/2} y$  is co-coercive<sup>9</sup> with constant 1, where  $D_s = (D + D^\top)/2$  and  $D_s^{-1/2}$  denotes the principal square root of the symmetric positive definite matrix  $D_s^{-1}$  and is therefore symmetric positive definite. Let us rename  $L := D_s^{-1/2}$  and  $Ly = \begin{bmatrix} v \\ w \end{bmatrix}$  and simplify the expression of  $G(y)$

$$\begin{aligned} G(y) &= LT(Ly) - L(D - D_s)Ly \\ &= L \left( \begin{bmatrix} F(v) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & A^\top \\ -A & 0 \end{bmatrix} Ly + \begin{bmatrix} 0 \\ b \end{bmatrix} \right) - L \begin{bmatrix} 0 & A^\top \\ -A & 0 \end{bmatrix} Ly \\ &= L \left( \begin{bmatrix} F(v) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \right). \end{aligned} \quad (46)$$

We now prove that the operator  $G(y)$  is co-coercive with constant 1, i.e.

$$(y_1 - y_2)^\top (G(y_1) - G(y_2)) - \|G(y_1) - G(y_2)\|^2 \geq 0. \quad (47)$$

Let us substitute (46) in the left-hand side of (47)

$$\begin{aligned}
& (y_1 - y_2)^\top (G(y_1) - G(y_2)) - \|G(y_1) - G(y_2)\|^2 \\
&= (y_1 - y_2)^\top (L \begin{bmatrix} F(v_1) \\ 0 \end{bmatrix} - L \begin{bmatrix} F(v_2) \\ 0 \end{bmatrix}) - \|L \begin{bmatrix} F(v_1) \\ 0 \end{bmatrix} - L \begin{bmatrix} F(v_2) \\ 0 \end{bmatrix}\|^2 \\
&= (Ly_1 - Ly_2)^\top (\begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}) - \|L \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}\|^2 \\
&= (\begin{bmatrix} v_1 - v_2 \\ w_1 - w_2 \end{bmatrix})^\top (\begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}) - \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix}^\top L^2 \begin{bmatrix} F(v_1) - F(v_2) \\ 0 \end{bmatrix} \\
&= (F(v_1) - F(v_2))^\top (v_1 - v_2) - [L^2]_{11} (F(v_1) - F(v_2)) \\
&\geq \alpha \|v_1 - v_2\|^2 - \|[L^2]_{11}\| \|F(v_1) - F(v_2)\|^2 \\
&\geq (\alpha - \|[L^2]_{11}\| \|L_F^2\|) \|v_1 - v_2\|^2 =: K \|v_1 - v_2\|^2,
\end{aligned}$$

The proof is concluded if  $K \geq 0$ . Let us compute  $[L^2]_{11} = [D_s^{-1}]_{11}$ . By inverting the block matrix  $D_s$  we get

$$[L^2]_{11} = \tau(I - \tau^2 A^\top A)^{-1} \succ 0. \quad (48)$$

Since  $\tau^2 A^\top A$  is symmetric positive semidefinite,  $\lambda_{\max}(\tau^2 A^\top A) = \tau^2 \|A\|^2 < 1$  by (24a) and  $\rho(\tau^2 A^\top A) < 1$ , i.e. the matrix is convergent. Hence, the Neumann series  $\sum_{k=0}^{\infty} (\tau^2 A^\top A)^k$  converges to  $(I - \tau^2 A^\top A)^{-1}$ . Substituting in (48) yields  $[L^2]_{11} = \tau \sum_{k=0}^{\infty} (\tau^2 A^\top A)^k \succeq 0$  and

$$\|[L^2]_{11}\| \leq \tau \sum_{k=0}^{\infty} (\tau^2 \|A\|^2)^k = \frac{\tau}{1 - \tau^2 \|A\|^2},$$

where we used the fact that the geometric series converges since  $\tau^2 \|A\|^2 < 1$  by (24a). Therefore  $K \geq \alpha - \frac{\tau}{1 - \tau^2 \|A\|^2} L_F^2$ . By condition (24b) we get  $\alpha \tau^2 \|A\|^2 + \tau L_F^2 < \alpha$  and thus

$$K \geq \frac{\alpha - \alpha \tau^2 \|A\|^2 - \tau L_F^2}{1 - \tau^2 \|A\|^2} > 0. \quad \blacksquare$$

**Lemma 5.** For all  $M \in \mathbb{N}$  it holds

$$\min_{y \in [0,1]^M} \lambda_{\min}(y \mathbb{1}_M^\top + \mathbb{1}_M y^\top) \geq -\frac{M}{4}. \quad (49)$$

*Proof:* The statement is trivially true for  $M = 1$ . For  $M > 1$ , problem (49) is equivalent to

$$\min_{\substack{y \in [0,1]^M \\ \|v\|=1}} v^\top (y \mathbb{1}_M^\top + \mathbb{1}_M y^\top) v = \min_{\substack{y \in [0,1]^M \\ \|v\|=1}} 2(v^\top y)(\mathbb{1}_M^\top v). \quad (50)$$

We show that (50) is negative by denoting  $\hat{y} = e_1$ ,  $\hat{v} = 0.6e_1 - 0.8e_2$  and observing that  $2(\hat{v}^\top \hat{y})(\mathbb{1}_M^\top \hat{v}) = -0.24$ . Let us consider a pair  $y^*, v^*$  minimizing (50) and note that  $\mathbb{1}_M^\top v^* \neq 0$ , because (50) is negative. We are left with two cases,  $\mathbb{1}_M^\top v^* > 0$  and  $\mathbb{1}_M^\top v^* < 0$ . Let us start analyzing  $\mathbb{1}_M^\top v^* > 0$ . To minimize  $2(v^\top y)(\mathbb{1}_M^\top v)$ , it must be

$$y_i^* = \begin{cases} 0 & \text{if } v_i^* > 0 \\ 1 & \text{if } v_i^* < 0, \end{cases} \text{ for all } i \in \{1, \dots, M\}. \quad (51)$$

Without loss of generality, we can assume  $y_i^* \in \{0, 1\}$  if  $v_i^* = 0$ . Hence we conclude that  $y^* \in \{0, 1\}^M$  and (49) reduces to

$$\min_{p \in \{0, \dots, M\}} \lambda_{\min} \left[ \begin{array}{c|c} 2(\mathbb{1}_p \mathbb{1}_p^\top) & \mathbb{1}_p \mathbb{1}_{(M-p)}^\top \\ \hline \mathbb{1}_{(M-p)} \mathbb{1}_p^\top & \mathbb{0}_{(M-p)} \mathbb{0}_{(M-p)}^\top \end{array} \right], \quad (52)$$

where without loss of generality we assumed the first  $p$  components of  $y^*$  to be 1 and the remaining to be 0. Note that the matrix in (52) features  $p$  identical rows followed by  $M - p$  other identical rows. Hence any of its eigenvectors must have  $p$  identical components followed by  $M - p$  other

identical components. With this observation and the definition of eigenvalue, it is easy to show that the matrix in (52) has only two distinct eigenvalues, the minimum of the two being  $p - \sqrt{Mp}$ . The function  $p - \sqrt{Mp}$  is minimized over the reals for  $p = M/4$  with corresponding minimum  $\lambda_{\min} = -M/4$ , as it can be seen by using the change of variables  $p = q^2$  and minimizing the quadratic function  $q^2 - \sqrt{M}q$ . Since  $p \in \{0, \dots, M\}$  in (52), the value  $-M/4$  is a lower bound for the minimum eigenvalue, and it is attained only if  $M$  is a multiple of 4. We conclude by noting that the derivation for the case  $\mathbb{1}_M^\top v^* < 0$  is identical to the derivation for the case  $\mathbb{1}_M^\top v^* > 0$  just shown, upon switching 0 and 1 in (51).  $\blacksquare$

*Proof of Proposition 4*

The constraints in (27), (28) can be expressed as  $\Gamma x \leq \gamma$  with

$$\Gamma = \begin{bmatrix} I_{M \cdot n} \\ -I_{M \cdot n} \\ -I_M \otimes \mathbb{1}_n^\top \\ \mathbb{1}_M^\top \otimes I_n \end{bmatrix}, \quad \gamma = \begin{bmatrix} \tilde{x} \\ 0 \\ -\theta \\ MK \end{bmatrix},$$

where  $\theta = [\theta^1, \dots, \theta^M]^\top$ , and  $\tilde{x} = [[\tilde{x}_t^i]_{i=1}^n]_{t=1}^M$ . Let us partition the constraint matrix  $\Gamma$  into its individual part  $\Gamma_1$  and coupling part  $\Gamma_2$

$$\Gamma = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} I_{M \cdot n} \\ -I_{M \cdot n} \\ -I_M \otimes \mathbb{1}_n^\top \end{bmatrix}, \quad \Gamma_2 = [\mathbb{1}_M^\top \otimes I_n] \quad (53)$$

and  $\gamma = [\gamma_1^\top, \gamma_2^\top]^\top$  accordingly. The KKT conditions for VI( $\mathcal{Q}$ ,  $F_N$ ) at the primal solution  $\bar{x}_N$  are [11, Proposition 1.3.4]

$$F_N(\bar{x}_N) + \Gamma_1^\top \mu + \Gamma_2^\top \lambda = 0, \quad (54a)$$

$$0 \leq \mu \perp \gamma_1 - \Gamma_1 \bar{x}_N \geq 0, \quad (54b)$$

$$0 \leq \lambda \perp \gamma_2 - \Gamma_2 \bar{x}_N \geq 0. \quad (54c)$$

Define  $\tilde{\mu}$  and  $\tilde{\lambda}$  as the dual variables corresponding to the active constraints (the other dual variables must be zero due to (54b) and (54c)). The KKT system (54) in  $\tilde{\mu}, \tilde{\lambda}$  only reads

$$\begin{aligned}
& \tilde{\Gamma}_1^\top \tilde{\mu} + \tilde{\Gamma}_2^\top \tilde{\lambda} = -F_N(\bar{x}_N), \\
& \tilde{\mu}, \tilde{\lambda} \geq 0,
\end{aligned} \quad (55)$$

where  $\tilde{\Gamma}_1, \tilde{\Gamma}_2$  contain the subset of rows of  $\Gamma_1, \Gamma_2$  corresponding to active constraints. To conclude the proof we need to show that (55) has a unique solution  $\tilde{\lambda}$ . To this end we apply the subsequent Lemma 6. To verify its assumption, we note that its negation is equivalent, given the expressions of  $\tilde{\Gamma}_1, \tilde{\Gamma}_2$  in (53), to the existence of  $R' \subseteq R^{\text{right}}$  such that for each vehicle  $i$  it holds  $\bar{x}_{N,t}^i \in \{0, \tilde{x}_t^i\}$  for all  $t \in R'$  or  $\bar{x}_{N,t}^i \in \{0, \tilde{x}_t^i\}$  for  $t \in \{1, \dots, n\} \setminus R'$  and such  $R'$  cannot exist by assumption.  $\blacksquare$

**Lemma 6.** Consider  $A_1 \in \mathbb{R}^{m \times n_1}$ ,  $A_2 \in \mathbb{R}^{m \times n_2}$ ,  $b \in \mathbb{R}^m$ . If the implication  $A_1 x_1 + A_2 x_2 = 0 \Rightarrow x_1 = 0$  holds, then the linear system of equations  $A_1 x_1 + A_2 x_2 = b$  has at most one solution in  $x_1$ .

*Proof:* Assume  $A\tilde{x} = b$  and  $A\hat{x} = b$ , then  $A_1 \tilde{x}_1 + A_2 \tilde{x}_2 = b$  and  $A_1 \hat{x}_1 + A_2 \hat{x}_2 = b$  imply  $A_1(\hat{x}_1 - \tilde{x}_1) + A_2(\hat{x}_2 - \tilde{x}_2) = 0$ , which by assumption implies  $\hat{x}_1 = \tilde{x}_1$ .  $\blacksquare$



## REFERENCES

- [1] M. K. Jensen, "Aggregative games and best-reply potentials," *Economic theory*, vol. 43, no. 1, pp. 45–66, 2010.
- [2] M. J. Smith, "The existence, uniqueness and stability of traffic equilibria," *Transportation Research Part B: Methodological*, vol. 13, no. 4, pp. 295–304, 1979.
- [3] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, no. 6, pp. 659–670, 2002.
- [4] H. Chen, Y. Li, R. H. Louie, and B. Vucetic, "Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1744–1754, 2014.
- [5] A. Cournot, *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*. Hachette, 1838.
- [6] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, "Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints," in *Proceedings of the IEEE Conference on Decision and Control*, 2016, pp. 6123–6128.
- [7] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton University Press, 2007.
- [8] A. Haurie and P. Marcotte, "On the relationship between Nash-Cournot and Wardrop equilibria," *Networks*, vol. 15, no. 3, pp. 295–308, 1985.
- [9] J. G. Wardrop, "Some theoretical aspects of road traffic research," *Proceedings of the institution of civil engineers*, vol. 1, no. 3, pp. 325–362, 1952.
- [10] Z. Ma, D. S. Callaway, and I. A. Hiskens, "Decentralized charging control of large populations of plug-in electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 67–78, 2013.
- [11] F. Facchinei and J. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.
- [12] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
- [13] M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.
- [14] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained deterministic mean field control," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2016.
- [15] F. Parise, B. Gentile, S. Grammatico, and J. Lygeros, "Network aggregative games: Distributed convergence to Nash equilibria," in *Proceedings of the IEEE Conference on Decision and Control*, 2015, pp. 2295–2300.
- [16] D. Paccagnan, M. Kamgarpour, and J. Lygeros, "On aggregative and mean field games with applications to electricity markets," in *Proceedings of the IEEE European Control Conference*, 2016, pp. 196–201.
- [17] J. Koshal, A. Nedić, and U. V. Shanbhag, "A gossip algorithm for aggregative games on graphs," in *Proceedings of the IEEE Conference on Decision and Control*, 2012, pp. 4840–4845.
- [18] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *4OR*, vol. 5, no. 3, pp. 173–210, 2007.
- [19] S. Kar and G. Hug, "Distributed robust economic dispatch in power systems: A consensus + innovations approach," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.
- [20] Y. Pan and L. Pavel, "Games with coupled propagated constraints in optical networks with multi-link topologies," *Automatica*, vol. 45, no. 4, pp. 871–880, 2009.
- [21] J. Koshal, A. Nedić, and U. V. Shanbhag, "Multiuser optimization: Distributed algorithms and error analysis," *SIAM Journal on Optimization*, vol. 21, no. 3, pp. 1046–1081, 2011.
- [22] F. Nemry, M. Brons *et al.*, "Plug-in hybrid and battery electric vehicles. Market penetration scenarios of electric drive vehicles," Institute for Prospective and Technological Studies, Tech. Rep., 2010.
- [23] J. Garcia-Villalobos, I. Zamora, J. San Martín, F. Asensio, and V. Aperribay, "Plug-in electric vehicles in electric distribution networks: A review of smart charging approaches," *Renewable and Sustainable Energy Reviews*, vol. 38, pp. 717–731, 2014.
- [24] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 940–951, 2013.
- [25] R. Arnott and K. Small, "The economics of traffic congestion," *American scientist*, vol. 82, no. 5, pp. 446–455, 1994.
- [26] S. Dafermos, "Traffic equilibrium and variational inequalities," *Transportation science*, vol. 14, no. 1, pp. 42–54, 1980.
- [27] T. Brinkhoff, "A framework for generating network-based moving objects," *GeoInformatica*, vol. 6, no. 2, pp. 153–180, 2002.
- [28] A. Sandmo, "Optimal taxation in the presence of externalities," *The Swedish Journal of Economics*, pp. 86–98, 1975.
- [29] F. Facchinei, A. Fischer, and V. Piccialli, "On generalized Nash games and variational inequalities," *Operations Research Letters*, vol. 35, no. 2, pp. 159–164, 2007.
- [30] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation*. Upper Saddle River, NJ: Prentice-Hall, 1986.
- [31] A. Nagurny, *Network economics: A variational inequality approach*. Springer Science & Business Media, 2013, vol. 10.
- [32] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, pp. 124–143, 1996.
- [33] P. Dubey, O. Haimanko, and A. Zapechelnyuk, "Strategic complements and substitutes, and potential games," *Games and Economic Behavior*, vol. 54, no. 1, pp. 77–94, 2006.
- [34] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 6, pp. 1393–1407, 2009.
- [35] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [36] G. Scutari, D. P. Palomar, F. Facchinei, and J. Pang, "Monotone games for cognitive radio systems," in *Distributed Decision Making and Control*. Springer, 2012, pp. 83–112.
- [37] S. Grammatico, "Dynamic control of agents playing aggregative games," *ArXiv e-prints*, Sep. 2016.
- [38] J. Pang, G. Scutari, D. P. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3251–3271, 2010.
- [39] L. Pavel, "An extension of duality to a game-theoretic framework," *Automatica*, vol. 43, no. 2, pp. 226–237, 2007.
- [40] P. Tseng, "Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming," *Mathematical Programming*, vol. 48, no. 1, pp. 249–263, 1990.
- [41] D. Bauso and R. Pesenti, "Mean field linear quadratic games with set up costs," *Dynamic Games and Applications*, vol. 3, no. 1, pp. 89–104, 2013.
- [42] H. Yin, U. V. Shanbhag, and P. G. Mehta, "Nash equilibrium problems with scaled congestion costs and shared constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1702–1708, 2011.
- [43] G. Wachsmuth, "On LICQ and the uniqueness of Lagrange multipliers," *Operations Research Letters*, vol. 41, no. 1, pp. 78–80, 2013.
- [44] T. K. Kristoffersen, K. Capion, and P. Meibom, "Optimal charging of electric drive vehicles in a market environment," *Applied Energy*, vol. 88, no. 5, pp. 1940–1948, 2011.
- [45] J. R. Correa and N. E. Stier-Moses, "Wardrop equilibria," *Wiley encyclopedia of operations research and management science*, 2011.
- [46] A. De Palma and N. Picard, "Route choice decision under travel time uncertainty," *Transportation Research Part A: Policy and Practice*, vol. 39, no. 4, pp. 295–324, 2005.
- [47] J. Li and H. M. Zhang, "Fundamental diagram of traffic flow : New identification scheme and further evidence from empirical data," *Journal of the Transportation Research Board*, 2011.
- [48] S. Schaible *et al.*, "Generalized monotone nonsmooth maps," *Journal of Convex Analysis*, vol. 3, pp. 195–206, 1996.